Jadavpur Journal of Languages and Linguistics
ISSN: 2581-494X

Volume 08, Issue 01

Editorial Board:
Atanu Saha
Indranil Dutta
Samir Karmakar

Managing Editor:
Samir Karmakar
Editorial Assistant:
Aniruddha Poddar

TABLE OF CONTENTS

Investigating the Origins of Sesquisyllabic Clusters in Proto-Tai: An Optimality Theory Analysis 001 – 011
Indrani Gogoi

Comparative Phonology of Multani and Majhi Punjabi: A Sociolinguistic Study 012 – 036
Isha Bhola

Phonological and Morpho-Syntactic Development in Early Childhood: A Cross-Sectional Analysis of Jordanian Arabic 037 – 046
Hosam Shatnawi & Mamta

A Comparative Study of Compound Word Formation in Assamese and Bodo languages 047 – 062
Banani Das

Derivational Adverbs by adding Locative Case Marker '-aw' in Bodo 063 – 069
Gojen Swargiary & Biswajit Brahma

Affixations in Kokborok 070 – 076
Priya Debbarma

Case Syncretism in Assamese through Caha's Framework: A Synchronic Analysis 077 – 088
Dewashree Ojah & Arup Kumar Nath

Multiple functions of -k(a) morpheme in early Assamese 089 – 100
Gitanjali Bez

Personal pronouns in Kokborok and Dimasa 101 – 111
Samir Debbarma & Kh. Dibyajyoti Barman

Individuative and Generic Referential NPs in Tutsa 112 – 124
Monjushna Sudhakaran & Gautam K. Borah

একক পদ ও একক বাচ্য 125 – 131
পিনাকী সরকার

Reflection on Linguistic Reality through Bhartṛhari's Vākyapadīya 132 – 138
Shubham Kumar

Navigating Linguistic Shifts: An Empirical Exploration of Internet-Driven Semantic Changes in College Students 139 – 157
Varshita Dhanda

Code-Mixing in Khasi-English: Patterns, Functions, and Sociolinguistic Implications 158 – 176
Ishita Chowdhury

The Glorious Spiritual Memory and Musical Space: The Women and the Language Matrix of Bhakti Songs in Challenging the Oppressive Hierarchies 177 – 200
Aditi Bhunia Chakraborty

Intergenerational Transmission of Tamil in Mauritius: A Sociolinguistic Analysis 201 – 210
Dahini Ponnen & S. Arulmozi

Cognitive Demands of Reading Devanagari Script in Formal Settings and Classrooms: A psycholinguistic perspective 211 – 227
Rupesh Kumar Pandey & Tariq Khan

Jadavpur Journal of Languages and Linguistics
ISSN: 2581-494X

Investigating the Origins of Sesquisyllabic Clusters in Proto-Tai: An Optimality Theory Analysis

Indrani Gogoi[1]
EFLU Hyderabad, Telangana

ARTICLE INFO

Article history:
Received 15/09/2025
Accepted 06/10/2025

Keywords:
Sesquisyllabic cluster;
Proto-Tai;
prefix;
Khmer;
Optimality Theory

ABSTRACT

The structure of the Proto-Tai prosodic word allows for it to exist either as a monosyllable or as a sesquisyllable. The sesquisyllabic cluster refers to words that consist of an unstressed "minor" syllable followed by a fully stressed "major" syllable. This paper deals with a set of minor syllables that may have originated as prefixes in Proto-Tai and later condensed into clusters in modern languages. This study aims to explore the origins of the [*k.σ_{maj}[2]] sesquisyllabic clusters drawing parallels between velar prefixes in Khmer and Thai. The various examples from Thai and other languages prove that the [*k] minor syllable in Proto-Tai is actually a velar prefix which indicates "generalized intensification". Additionally, an Optimality Theory (OT) approach is employed to examine the phonotactic constraints influencing syllable formation. This study contributes to a broader understanding of Tai historical phonology and the mechanisms underlying syllable evolution in the language family.

Introduction

The structure of the Proto-Tai prosodic word allows for it to exist either as a monosyllable or as a sesquisyllable. Proto-Tai is the ancestor of the Tai languages, the largest branch of the Tai-Kadai language family, of Mainland Southeast Asia, which includes modern Lao, Shan, Tai Lü, Tai Dam, Ahom, Northern Thai, Standard Thai, Bouyei, and Zhuang. Since no written records of Proto-Tai exist, linguists have reconstructed it using the comparative method. Notably, Li Fang-Kuei reconstructed the language in 1977, followed by Pittayawat Pittayaporn in 2009. The Tai language Proto-Tai has a rich array of complex onsets in addition to simple onsets, which can be either tautosyllabic or sesquisyllabic. The term "sesquisyllable", coined by Matisoff (1973), refers to words that consist of an unstressed "minor" syllable followed by a fully stressed "major" syllable. This paper deals with a set of minor syllables that may have originated as prefixes in Proto-Tai and later condensed into clusters in modern languages. The idea stems from the fact that sesquisyllabic clusters do not conform to the Sonority Sequencing Principle (SSP), justifying the existence of these clusters (explained in detail later). However, some sesquisyllabic clusters deviate from this pattern, i.e they adhere to the SSP, yet they still exhibit characteristics typical of sesquisyllabic structures. This study focuses on the [*k] prefix in Proto-Tai and analyzes it using Optimality Theory. The paper is divided into seven sections, with the introduction as the first. The second section deals with the research methodology followed by the

[1] Corresponding Author Email ID: indranigogoi3@gmail.com
[2] σ_{maj} = Major syllable

syllable structure of the proto language. The fourth section gives a description of the prefixes in Proto-Tai and the fifth section deals with the /k-/ prefix in particular. All the findings are accounted for using Optimality-theoretic constraints in section six and section seven puts forth a summary and a conclusion.

Research Methodology

Since this paper deals with the Proto-Tai language, all the data used to study are taken from secondary sources. It has been taken from:

(i) Previous research works

Most of the data has been gathered from The Phonology of Proto Tai by Pittayaporn (2009).

(ii) Online dictionaries

An online dictionary titled "Thai Lexicography Resources." was utilized for collecting Proto-Tai words.

(iii) Native speaker's knowledge

Certain data was obtained through discussions with native speakers, who provided insights and confirmation regarding the language in question. This information was further validated through conversations with community members proficient in the language. It is essential to understand the syllable structure of this proto language to better understand how these sesquisyllables differ from others.

Syllable Structure in Proto-Tai

The conventional belief regarding the syllable structure of Proto-Tai suggests it was exclusively monosyllabic. Despite not explicitly stating this in the Handbook of Comparative Tai (Li 1977), Li's reconstructed forms are all monosyllabic, a view shared by other linguists working on Proti-Tai reconstruction like Nishida (1954; 1955), Liang and Zhang (1996), Luo (1997), Sarawit (1973), and Gedney (1972a, 1989b). This monosyllabic perspective likely stems from the prevalence of monosyllabic cognates in modern Tai varieties, a factor that significantly influences different authors' segmental reconstructions.

Recent advancements in Kra-Dai language studies challenge the notion of Proto-Kra-Dai being exclusively monosyllabic. An alternative to this monosyllabic perspective is considering Proto-Tai as a sesquisyllabic language (Pittayaporn, 2009). Edmondson and Yang (1988), as well as Ferlus (1989; 1996b), also support the sesquisyllabic hypothesis through initial correspondences among Kam-Sui languages, despite these modern languages being monosyllabic. Sesquisyllabic structures differ from tautosyllabic clusters in this structural definition where the consonants in a sequence at the start of a sesquisyllable are not all associated with a single syllable node.

To sum up, these assumptions vary from monosyllabic (as conventionally suggested by Li in 1977) to sesquisyllabic (proposed by Ferlus in 1990) or disyllabic (as argued by Haudricourt in 1975 and Strecker in 1983). This paper will consider Proto-Tai as a sesquisyllabic language. Every PT prosodic word was distinctly characterized by a single lexical tone.

	Open syllable		Closed syllable
Monosyllable	*C(C)(C)V^{T}	C(C)(C)VV^{T}	C(C)(C)V(V)C^{T}
Sesquisyllable	*C(C).C(C)(C)V^{T}	C(C).(C)(C)(C)VV^{T}	C(C).C(C)(C)V(V)C^{T}

*C = consonant, V = vowel, T = tone, * = not a possible word*

Table 1: *Canonical shape of Proto-Tai prosodic word (Pittayaporn; 2009)*

The Proto-Tai prosodic word can be segmented into three parts: 1) onset, 2) rime, and 3) tone. The onset comprises any consonant that precedes the vocalic nucleus, encompassing both the tautosyllabic and sesquisyllabic onsets, as defined earlier. The rime includes the vocalic nucleus and, if present, the final consonant. Finally, the tone represents the suprasegmental contrast associated with the major syllable, while the minor syllable does not carry any tone.

They can be either open or closed syllables, as illustrated in Table 1. However, open syllables with a short vowel are prohibited in this language. This implies that only a heavy syllable could independently function as a prosodic word in the case of monosyllables or act as the major syllable in sesquisyllables.

Complex onsets in Proto-Tai

As already mentioned, Proto-Tai has both tautosyllabic as well as sesquisyllabic clusters. In a monosyllabic word, it comprises an onset and a heavy rime. For a sesquisyllabic word, the major syllable comprises an onset and a heavy rime.

Consonant clusters in Proto-Tai are created through stops combining with the medial *-r and *-l- resulting in clusters like *pr-, *pl-, *kr-, *kl-, *bl- and so on. Additionally, some consonants also form clusters by combining with the medial *-w- as seen in *pw-, *tw-, *kw- and so on. These clusters in Proto-Tai can be categorized into labial, dental, velar, and uvular clusters. All these clusters adhere to the Sonority Sequencing Principle, which states that "in any syllable there is a segment constituting a sonority peak that is preceded and/or followed by a sequence of segments with progressively decreasing sonority values (Selkirk 1984:116)." This principle mandates that onsets follow a rising sonority pattern, whereas codas follow a declining sonority pattern. The following sonority scale has been adopted for this study that arranges segments from more sonorous to less sonorous.

0= voiceless stops	4= nasals
1= voiced stops	5= liquids
2 = voiceless fricative	6= glides
3 = voiced fricatives	7 = vowels

Table 2: *Sonority Hierarchy*

In the list given above, voiceless stops are the least sonorous and can be said to have the sonority value of 0. On the opposite end, vowels are the most sonorous with the sonority value of 7. If the consonant sequence obeys the SSP, it may be realized as a tautosyllabic cluster.

Proto-Tai also comprises various sesquisyllabic clusters. Sesquisyllabic clusters can have up to two consonants in the minor syllable and up to three consonants in the major syllable. (Pittayaporn: 162). The sesquisyllabic clusters found in Proto-Tai are:

a) Voiceless stop + voiceless stop (*p.t-, *k.t and *p.q-)
b) Voiceless obstruent + voiced stop (*Ç.b, *Ç.d)
c) Voiced obstruent + voiceless stop (*Ç.t, *Ç.k)
d) Voiceless stop + liquid/glide (*k.r-, *p.r- and *Ç.w)
e) Voiced consonant + liquid/glide (m.l, *Ç.r, *Ç.l)

All the sesquisyllabic clusters mentioned above violate the Sonority Sequencing Principle which accounts for their formation in this language. However, an exception arises in clusters where liquids or glides follow voiceless consonants (d and e), as these do not violate the SSP. One possible explanation for this is a theory that minor syllables may have originally been prefixes which gradually shortened over time and eventually transformed into single consonants or clusters within a syllable. Before delving into the /k-/ prefix, let us first understand the prefixes in Proto-Tai.

Prefixes in Proto-Tai

In Proto-Tai, most words or morphemes are monosyllabic. However, many polysyllabic words also exist, primarily as compounds and loanwords from Indic languages such as Sanskrit and Pali, as well as from Khmer, an Austroasiatic language spoken mainly in Cambodia.

There are no inflections in Tai similar to -ed in English ‘walked’ or -s in ‘dogs.’ Instead, new words are primarily formed through compounding, such as nâa-taa ‘countenance’ (literally, ‘face-eye’), kèp-kìaw ‘to harvest’ (literally, ‘gather-cut with a sickle’). Reduplication, the repetition of a word or part of it, is also common. For example, díi-dii ‘very good’ from dii ‘good.’ Partial reduplication is also found, as seen in sanùk-sanăan ‘to enjoy oneself’ derived from sanùk ‘to have fun.’

Some Tai languages have developed prefixes by shortening words that were once full words. Additionally, prefixes and infixes are also common in the Indic and Khmer loans into the more southerly Tai languages. For example, pra-thêet ‘country’ (from Sanskrit pradeśa) and d-amn-ɤɤn ‘to proceed,’ derived from dɤɤn ‘walk’ tracing back to the Khmer damnaɤ (‘gait, bearing’) and dae (‘walk’). There also emerged some native prefixes that are shortened forms of words that were once full words. For example, the prefix ma-, found in fruit names like ma-phráaw (‘coconut’) and ma-mûaŋ (‘mango’), originates from màak (‘areca nut’), which originally meant ‘fruit.’ Similarly, in sa-dɯˇɯ (‘navel’), sa- is a reduced form of săaj (‘line, string’), referring to the umbilical cord. Older derivational processes involved consonant or tone alternation, as seen in níi (‘this’) and nîi (‘here’), nɔɔj (‘small, little’) and nɔɔj (‘a little bit’), and khiaw (‘sickle’) and kìaw (‘to cut with a sickle’). However, these processes are no longer active. (Li and Solnit, 2023). Now, let us examine the k- prefix in Tai languages in greater detail.

/*k-/ prefix in Proto-Tai

The presence of the /k-/ prefix in languages such as Thai and Khmer has been acknowledged by earlier researchers. Jenner’s work (1982) demonstrated the existence of the /ka-kra/ prefixes in Khmer, which serve multiple functions like “a velar prefix which indicates “generalized intensification”, a reduplicated initial consonant, also indicating intensification; and a velar

prefix which designates nouns of five specific types." Below are some examples that demonstrate "generalized intensification" in Khmer.[3]

1) "Concentration of form or range of action, ŋooc 'to draw back' > kŋooc 'to flinch
2) Distinction of form or feature, wil 'to turn' > kwəl 'to turn rapidly'
3) Something negative, weək 'to be blocked' > kwak 'to be blind'"

Varasarin (1984) argues that velar prefixes in Thai operate similarly to those in Khmer. Moreover, Varasarin proposes that the variations observed in Thai can be traced back to Khmer influences, indicating a linguistic connection and potential derivation of certain linguistic features from Khmer to Thai.

Thai and its prefixes

Bilmes (1998) explored velar prefixes in Khmer that specifically designate various types of nouns related to semantic fields like human beings, other animate beings, and substances. Interestingly, Thai has introduced additional senses not found in Khmer. Bilmes (ibid) claims that Thai prefixes have been borrowed from Khmer, as among the Tai-Kadai languages, only Siamese Thai exhibits words following this pattern.

The prefixes which existed in Khmer and were later borrowed to Thai cover semantic fields and these fields together form a radial category. A radial category has a central member and peripheral members where the central member is considered to be the best in the category. Radial categories in Thai include animals, plants, substances and vessels and containers. Some examples from each category are discussed below:

➢ **Plants**

karaŋ	coral
kapraw	basil
kalɔ̌ɔ̌n	variety of mango

➢ **Plant parts and products**

kalaa	hard shell of a coconut
kataŋ	gum
kathi	coconut cream

➢ **Vessels/ containers made of plants**

kalaa	half a coconut used as a bowl
kradoŋ	wide, flat shallow basket for winnowing rice

Despite the radial categories discussed above, it is also seen that some other tokens do not fit into any of those categories. They are characterized by the property of "generalized intensification" (Bilmes 1998).

A similar scenario can be observed in the clusters in Proto-Tai.

*Proto-Tai and its /*k-/ prefix*

As already discussed, Proto-Tai has a rich array of sesquisyllabic clusters. They can form clusters using the various combinations like voiceless stop + voiceless stop, voiceless obstruent + voiced stop, voiced obstruent + voiceless stop, voiceless stop + liquid and voiced consonant + liquid.

Here, we propose a theory that *all the minor syllables in the sesquisyllabic clusters of Proto-Tai might have originated as prefixes that gradually condensed into shorter syllables and*

[3] The examples used are from Bilmes (1998).

eventually transformed into single consonants or clusters within a syllable. We will try to explain this with the help of the words beginning with the minor syllable /*k-/.

As discussed earlier, Khmer has a set of velar prefixes which indicates what Jenner and Pou called "generalized intensification." Like Khmer, Proto-Tai has the /*k-/ prefix which indicates generalized intensification discussed in the examples below:

a. *haŋ 'to trap' >*k.raŋ 'to imprison'
b. *t^{h}om 'to fill, cover, bury' >*k.tem 'full'
c. *ton 'origin' >*k.tɯn 'to wake up'
d. *rip 'to take by force >*k.rap 'to drive away'

It has been claimed that /k-/ words in Thai cover various semantic fields and that these fields come together to form a radial category. We saw earlier that /k-/ derivatives in Thai designate plants, plant parts and vessels made of plants. This can be applied to its Proto language as well, where we can see /k-/ derivatives like *k.toŋ 'big leaf'. It can be substantiated by the presence of the similar word in Thai - kra.t^{h}oŋ 'vessel made of banana leaves'.
The examples discussed above show that the [*k-] syllables in those words in Proto-Tai could be a prefix. This differentiates these words from words that have [*kr-] as a true onset cluster.[4]

An OT analysis
The syllable structure in Proto-Tai allows segments in the onset, nucleus and coda positions. We now attempt to provide an OT analysis for the same. The two markedness constraints that can be developed for Proto-Tai are as follows:

ONSET: Syllables must have onsets (Itô 1989, Prince and Smolensky 1993):

This constraint means that syllables cannot start with vowels. It only applies to syllables that begin with a consonant sound. So, in languages where this constraint is the most important, every syllable needs to start with a consonant sound.

NO-CODA: Syllables are open

This constraint states that syllables cannot end in a consonant. In languages where this constraint is higher ranked, every syllable will always end with a vowel, resulting in only open syllables.
The faithfulness constraints observed in Proto-Tai are as follows:

DEP-IO:"Output segments must have input correspondents." No epenthesis. (McCarthy & Prince 1995:16)

This constraint is critical for ensuring that the output remains consistent and relevant to the input data.

MAX-IO: Input segments must have output correspondents. No deletion

This constraint is critical for ensuring that the output remains consistent and relevant to the input data.
Therefore, the constraint hierarchy for Proto-Tai syllable structure is as follows:

[4]With regard to /*p-/ we assume these are borrowed words or they arose due to the influence of Austronesian languages where **p(a)*- was a prefix used for non-stative (i.e., dynamic) causatives (Blust 2009:282). It needs further investigation.

DEP-IO, MAX-IO >> ONSET >> NO-CODA

Since PT allows consonant clusters in the onset position, we can account for it using the constraint discussed below:

***COMPLEX^{ONS}**: Simple onsets only

This markedness constraint states that onsets must always be simple. Any language that has complex onsets violates this constraint and thus remains dominated. This constraint is ranked lower in Proto-Tai as this language allows consonant clusters in the onset position. On the other hand, the constraint

***COMPLEX^{CODA}**: Simple codas only

This constraint is high ranked as Proto-Tai disallows complex codas.

DEP-IO, MAX-IO, *COMPLEX^{CODA}>> ONSET >> NO-CODA, *COMPLEX^{ONS}

Apart from tautosyllabic clusters in the onset, Proto-Tai also has a rich array of sesquisyllabic clusters. This can be accounted for with the help of the following markedness constraint:

Sonority Sequencing Principle (SSP) (Selkirk 1984:116): "Onsets must rise in sonority while codas must fall in sonority."

This is a markedness constraint which states that sonority must be highest at the syllable peak and decrease towards the margins. This constraint is very high ranked in Proto-Tai as clusters violating SSP are not allowed in this language. To avoid such a cluster, the word is divided into two syllables- a minor syllable followed by a major syllable. The fact that Proto-Tai does not prefer epenthesis or deletion to avoid such a cluster shows that MAX-IO and DEP-IO dominate SSP.

Therefore,

DEP-IO, MAX-IO,*COMPLEX^{CODA}>> SSP >> ONSET >>NO-CODA, *COMPLEX^{ONS}

The ranking is illustrated by the following tableau:

/*pta:/	DEP-IO	MAX-IO	*COMPLEX^{CODA}	SSP	ONSET	NO-CODA	*COMPLEX^{ONS}
☞ a. p.ta:							
b. pta				*!			*
c. pa:		*!					
d. pa.ta:	*!						

Table 3: *Tableau for sesquisyllabic cluster /p.t-/*

In this tableau, we note that the consonant cluster /pt/ in the onset position would violate SSP as both the segments are of the same sonority if they are considered part of a single syllable. Hence, if they are made tautosyllabic, the constraint SSP is violated (as seen in candidate b.). Candidates c and d violate MAX-IO and DEP-IO respectively and are thus not optimal. The only way to make it optimal is by creating a sequisyllable, wherein the /p/ is not part of the major syllable but forms a minor syllable of its own. In most of these sesquisyllables, there would be a very short vowel inserted at the phonetic level for ease of articulation.

To sum up, we note that consonant clusters that obey SSP are tautosyllabic in Proto-Tai but those that do not have a minimum sonority distance are sesquisyllabic. The only exceptions

to these, we saw in the previous section, are the sequences where we have voiceless stops followed by liquids/glides. We hypothesized in these cases that the voiceless stops were in fact, prefixes rather than part of the stem. To account for the occurrence of these clusters in sesquisyllables, we use the constraint ALIGN-MORPH-L. If [k] and [p] seen in these sesquisyllables are prefixes, they would be separate morphemes prefixed to the stem.

ALIGN-MORPH-L (Kager 2004: 115)

The left edge of a morpheme coincides with the left edge of a syllable.

This constraint explains why such clusters go for the formation of sesquisyllables. The following tableau shows the ranking of the constraints. ALIGN-MORPH-L would also be a high ranked constraint.

The ranking is illustrated by the following tableau:

/*k+rap/ 'to drive away'	DEP-IO	MAX-IO	ALIGN MORPH-L	*COMPLEX^{CODA}	SSP	ONSET	NO-CODA	*COMPLEX^{ONS}
a. [krap]			*!				*	*
☞b. [k.rap]							*	
c. [kap]		*!					*	
d. [kapa]	*!	*						

Table 4: *Tableau for sesquisyllabic /k.rap/*

Candidate b is the optimal candidate as it violates only the low ranked constraint of NO-CODA. Candidate a violates ALIGN-MORPH-L as the word is made up of two morphemes, the prefix /k-/ and the stem /rap/. The left edge of the morpheme /k/ is aligned to the left edge of the syllable, but the left edge of the morpheme /rap/ is not aligned to the left edge of the syllable, if it is part of the onset. It also violates the lower ranked constraints NO-CODA and *COMPLEX^{ONS}. Candidates c and d are also sub-optimal. Hence, the only option in the language is to make it sesquisyllable.

Therefore, the constraints for Proto-Tai are ranked thus:

DEP-IO, MAX-IO >> ALIGN-MORPH–L >> SSP >> ONSET >> No-CODA, *COMPLEX^{ONS}

Now, let us see if this constraint ranking can account for the tautosyllabic clusters allowed in this language. The advantage of Optimality Theory is that once the ranking of the constraints is established, it should account for all phenomena in the language.

The ranking is illustrated by the following tableau:

/*krap/ (gloss unavailable)	DEP-IO	MAX-IO	ALIGN-MORPH-L	*COMPLEX^{CODA}	SSP	ONSET	NO-CODA	*COMPLEX^{ONS}
☞a. krap							*	*
b. kə.rap	*!						*	
c. kap		*!					*	

d. k.rap						*	

Table 5: *Tableau for tautosyllabic /krap/*

In this case, there is no violation of the constraint ALIGN-MORPH-L as the /k-/ is part of the stem and not a prefix and /-rap/ is not a separate morpheme. Hence, candidate a is the optimal one. Candidates b and c are sub-optimal as they violate DEP-IO and MAX-IO respectively. Candidate d. would require another constraint NO MINOR SYLLABLES, which needs to be high ranked to disallow the creation of sesquisyllables which should be lower ranked than ALIGN-MORPH-L.

NO MINOR SYLLABLES

A minor syllable is disallowed.

To summarize, Proto-Tai has complex onsets which obey the SSP. The clusters violating SSP form sesquisyllables. In some cases, the first consonant in a cluster is a prefix rather than part of the stem.

Therefore, the constraint hierarchy for syllable structure in Proto-Tai is as follows:

DEP-IO, MAX-IO >> *COMPLEX^{CODA} >> ALIGN-MORPH-L >> SSP >> ONSET >> NO-CODA, *COMPLEX^{ONS}

Summary and Conclusion

This paper investigates the origins of sesquisyllabic clusters in Proto-Tai, focusing on the /k-/ prefix. The investigation of the /k-/ prefix in Proto-Tai revealed similarities with Khmer, where velar prefixes serve semantic functions such as intensification and categorization. Comparative evidence from modern Thai further supports this hypothesis, as words with /k-/ derivatives in both languages exhibit shared semantic domains. This suggests that the minor syllable in sesquisyllabic clusters may have originated as a prefix rather than a true consonant cluster. The study also incorporates an Optimality Theory approach to analyze the constraints governing syllable structure in Proto-Tai.

Overall, this research contributes to the understanding of Tai historical phonology and the evolution of syllable structure in the language family. This supports the hypothesis that sesquisyllabic structures in Proto-Tai resulted from historical prefixation processes. The /k-/ prefix, in particular, played a crucial role in shaping the phonological development of Tai languages. Future studies could further investigate other prefixes and their impact on Tai phonology and morphology, shedding more light on the evolutionary pathways from Proto-Tai to its modern descendants.

References

Bilmes, L. (1998). The /ka-/ and /kra-/ prefixes in Thai. *Linguistics of the Tibeto-Burman Area*, 21 (2), 73-96.

Blust, R. A. (with Australian National University). (2009). *The Austronesian languages*. Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University.

Edmondson, J. A., & Yang, Q. (1988). Word-Initial Preconsonants and the History of Kam-Sui Resonant Initials and Tone. In J. Edmondson, & D. Solnit (Eds.), Comparative Kadai: Linguistic Studies beyond Tai (pp. 143-166).Dallas: Summer Institute of Linguistics and University of Texas at Arlington.

Ferlus, Michel. (1982). Spirantisation des obstruantes médiales et formation du système consonantique du vietnamien. Cahiers de linguistique Asie Orientale 11 (1): 83- 106.

Ferlus, Michel. (1989). Remarques sur le consonantisme des langues kadai. Paper presented at the 22nd International Conference on Sino-Tibetan Languages and Linguistics. University of Hawai'i, Manoa. October 5-8, 1989.

Gedney, W. J. (1972). A puzzle in comparative Tai phonology. *Tai phonetics and phonology*, 52-57.

Gedney, W. J. (1989). Future directions in Comparative Tai Linguistics. In R. J. Bickner, J. Hartmann, T. J. Hudak & P. Peyasantiwong (Eds.), Selected papers on Comparative Tai Studies (pp. 7-116). Ann Arbor: Center for Southeast Asian Studies, the University of Michigan.

Haudricourt, André-Georges. (1975). A propos du puzzle de W.J. Gedney. In J. G. Harris & J. R.Chamberlain (Eds.), Studies in Tai linguistics in honor of William J. Gedney (Vol. 252-258). Bangkok: Central Institute of English Language.

Ito, J. (1989). A prosodic theory of epenthesis. *Natural Language and Linguistic Theory*, *7*(2), 217–259. https://doi.org/10.1007/BF00138077

Jenner, Philip N & Saveros POU. (1982). A Lexicon of Khmer Morphology. *Mon-Khmer studies IX-X*. In Philip N. Jenner (Ed.). Hawaii: The University Press of Hawaii.

Kager,R.(1999). *Optimality Theory* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511812408

Li, F. K. (1977). A Handbook of Comparative Tai. *Oceanic Linguistics Special Publications*, *15*, i–389. http://www.jstor.org/stable/20006684

Li, F.K., Solnit, D.B. (2023, August 11). Tai languages. Encyclopedia Britannica. https://www.britannica.com/topic/Tai-languages

Liang, M. & Zhang, J. (1996) Dong-tai Group Conspectus. Beijing: Chinese Academic Social Science Press.

Luo, Yongxian. (1997). Expanding the Proto-Tai Lexicon- a Supplement to Li (1977). *The Mon-Khmer Studies Journal*, 27, 271-298.

Matisoff, James A. (1973). Tonogenesis in Southeast Asia. In L. M. Hyman (Ed.), Consonant types & tones (pp. 71-95). Los Angeles: The Linguistic Program, University of Southern California.

McCarthy, John and Alan Prince. (1993). Prosodic Morphology I: constraint interaction and satisfaction. Ms. University of Massachusetts, Amherst, and Rutgers University, New Brunswick, NJ

Nishida, Tatsuo. (1954). Tonēmu ni yoru Tai shogo hikaku gengogakuteki kenkyū [Comparative linguistic studies of the Tai languages based on tonemes]. Gengo Kenkyū, 25, 19-46.

Nishida, Tatsuo. (1955). Makku-Sui go to kyōtsu go hikaku gengogakuteki kenkyū [Mak-Sui and Common Tai]. Gengo Kenkyū, 28, 30-62.

Pittayaporn, P. (2009). *The Phonology of Proto-Tai*. [Doctoral dissertation, Cornell University].

Pittayaporn, P. (2009). Proto-Southwestern Tai revised. Journal of the Southeast Asian Linguistics Society, 2: 212-244.

Prince, A. S., & McCarthy, J. J. (2001). *Prosodic Morphology: Constraint Interaction and Satisfaction* [Application/pdf]. Rutgers University. https://doi.org/10.7282/T3B856GM

Prince, A., & Smolensky, P. (2004). Optimality Theory: Constraint Interaction in Generative Grammar. In J. J. McCarthy (Ed.), *Optimality Theory in Phonology* (1st ed., pp. 1–71). Wiley. https://doi.org/10.1002/9780470756171.ch1

Sarawit, Mary E. (1973). The Proto-Tai vowel system. Unpublished Ph.D. dissertation, University of Michigan, Ann Arbor. Selkirk, Elizabeth O. (1984). On the major class features and syllable theory. In M. Aronoff & R.T. Oehrle (Eds.), Language sound structure (Vol. 107-136). Cambridge, MA: MIT Press

Strecker, D. (1982). Proto-Tai vowels revisited: a companion and critique of the work of Sarawit and Li. Linguistics of the Tibeto-Burman Area, 7 (2), 33-74.

Varasarin, Uraisi. (1984). Les éléments khmers dans la formation de la langue siamoise. (Langues et civilisations de l'Asie du sud-est et du monde insulinien, 15.) Paris: SELAF. 411pp. (Summary in English, German, Spanish, and Thai Bibliography: p.[401]-406).

rflr.org - Linguistics - Tai-Kadai Family - Survey. (n.d.). https://www.rflr.org/linguistics/tai-kadai/survey/

SEAlang Library Phake Lexicography. (n.d.). http://sealang.net/phake/

Thai lexicography resources. (n.d.). http://sealang.net/crcl/proto/

Jadavpur Journal of Languages and Linguistics
ISSN: 2581-494X

Comparative Phonology of Multani and Majhi Punjabi: A Sociolinguistic Study

Isha Bhola[1]
University of Delhi, Delhi

ARTICLE INFO

Article history:
Received 15/09/2025
Accepted 06/10/2025

Keywords:
Punjabi language,
lexical alterations,
dialectal variations,
Multani Punjabi,
sociolinguistic factors,
phonological modifications

ABSTRACT

Significant changes are occurring in the Punjabi language, especially in the phonological and lexical domains. The effects of these alterations are emphasized by Arslan (2021) and Hasan (2021), with Arslan emphasizing the possible loss of distinguishing sounds and Hasan on the vocabulary variations between the Majhi and Dhani dialects. Social and political variables further compound these changes, as demonstrated by Alizai's (2021) study on the linguistic shift of Punjabi speakers in the province of Baluchistan. This paper aims to investigate Multani Punjabi's phonological modifications to the Punjabi Lexicon. Multani Punjabi is spoken in Faridabad, Ballabhgarh, Palwal, and a few more areas Jain, V. (2019). Purposive sampling of 20 speakers, 10 from each dialect will be used in the process to gather information on phonemic and phonological variances. Particularly, the Multani Punjabi dialect is losing tonality, which has an impact on the meaning variances of words that have tonality (Rafi, 2010). The intricate interaction of sociolinguistic and phonological elements in the development of the Punjabi vocabulary is reflected in these modifications. The purpose of this research is to clarify Multani Punjabi's changing phonological characteristics, the phonemic comparison between Multani and Mahji using the Levensthein algorithm, and their consequences in the larger context of language development through a qualitative examination of the data.

Introduction

The Indo-Aryan branch of the Indo-Iranian language family, which is a branch of the larger Indo-European language family, is where the Multani Punjabi language is found. One of the largest language families in the world, Indo-European is thought to have originated. Vedic Sanskrit, which was spoken in ancient India 1500 BCE, is the earliest known ancestor of Multani Punjabi within the Indo-Iranian branch. Over time, contact with other languages and regional dialects had an impact on the development of the Old Punjabi language. Punjabi had further been divided into various regional dialects by the Middle Ages, one of which was Multani Punjabi, which is spoken in the Multan area of modern-day Pakistan.

Like other Punjabi dialects, Multani Punjabi has its own unique vocabulary, phonetics, and grammar. It is quite similar to other Punjabi dialects, but it also has distinct characteristics that are shaped by the linguistic, cultural, and historical background of the area.

[1] Corresponding Author Email ID: ishabhola2310@gmail.com

Background of the Study

Linguistic groups' dynamic contacts and impacts are reflected in phonological changes in languages. Language contact scenarios give rise to loanwords, which are a prime example of how languages change by borrowing from other linguistic sources. Loanwords are lexical terms imported from one language into another, frequently due to necessity or prestige, according to Trudgill (1992) and Crystal (2008). Words from other languages are adopted by speakers due to the need to bridge lexical gaps and the desire for linguistic status (Mahmood, Hussain, & Mahmood, 2011).

With six major languages and fifty-nine minor languages, Pakistan is a linguistically diverse country (Rahman, 2006). Seventy-three languages are spoken in Pakistan, according to Simons and Fennig (2017), who documented them in the Ethnologue: Languages of the World. Among these languages, Urdu, English, Punjabi, Sindhi, Pashto, and Balochi are particularly well-known. English, an official language, and Urdu, the nation's language, are essential in the fields of law, journalism, and education. In the meantime, about 39% of Pakistan's population speaks Punjabi, a regional language that is the main language of the Punjab province (Population Census, 2017).

Investigating Multani Punjabi's phonological changes and their sources in the Punjabi vocabulary is the goal of this study. There have been noticeable tone changes in Multani Punjabi, which is spoken in places like Faridabad, Ballabhgarh, and Palwal. These changes have affected the subtle meanings of words that are sensitive to tonal shifts (Rafi, 2010). This study will clarify the changing phonological traits of Multani Punjabi and their wider consequences in the context of language development by utilizing selective sampling and qualitative methods like interviews and audio recordings.

The phonemic inventory of the Multani Punjabi:

Consonants

	Bilabial	**Dental**	**Alveolar**	**Retroflex**	**Palatal**	**Velar**	**Glottal**
Plosive	p, b		t, d	ʈ, ɖ		k, g	
Nasal	m		n			ŋ	
Trill			r				
Tap/Flap							
Fricative		f, v	s, z		ʃ	x	h
Lateral Fricative			ɬ				
Approximant			ɹ		j	w	

Lateral Approximant	l

Vowels

Front	**Central**	**Back**
i		u
e		o
ɛ	ə	ɔ
	a	

Statement of the Problem

There are many major phonological changes occurring to Punjabi, especially in the Multani dialect that is spoken in Palwal, Ballabhgarh, Faridabad, and the adjacent districts. The Punjabi vocabulary is affected by these modifications, which are caused alterations to lexical and semantic elements. Comprehending and recording these phonological changes in Multani Punjabi is crucial in order to appreciate the wider consequences for language evolution and maintenance in the face of sociolinguistic influences. The goal of this study is to ascertain the phonological changes, the phonemic comparison, cause and effects of the phonological alterations that have been noted in Multani Punjabi.

The study's special objective, which is to evaluate and explain Multani Punjabi's evolving phonological characteristics and the phonemic comparison within the framework of broader language development, sociolinguistic dynamics, and the preservation of linguistic legacy, is highlighted in this problem statement.

Aim and Objectives

This study delves into the intricate phonological changes occurring in Multani Punjabi within the framework of sociolinguistics, aiming to discern the influential factors shaping these changes and their implications for the evolution and unique identity of this dialect within the Punjabi language family. The research methodology involves a multifaceted approach, starting with a phonemic comparison between Multani Punjabi and the standard Majhi dialect using the Levenshtein algorithm. This quantitative analysis will provide insights into the specific phonological variations distinguishing Multani Punjabi from the established standard. Additionally, the study will conduct an extensive review of existing literature to contextualize these observed phonological alterations within the broader linguistic landscape. By synthesizing findings from previous studies, the research aims to establish a comprehensive understanding of the historical and sociocultural underpinnings influencing the phonetic evolution of Multani Punjabi. Empirical data collection will be carried out through purposive sampling of Multani Punjabi speakers in targeted regions such as Faridabad, Ballabhgarh, Palwal, and adjacent areas. Qualitative research techniques including audio recordings and structured interviews will be employed to document and analyze spontaneous speech patterns and phonological features unique to Multani Punjabi. Through this approach, the study seeks to capture the nuanced

phonetic nuances that contribute to the dialect's distinctiveness and resilience. The ultimate goal of this research is to pinpoint and record significant phonological changes taking place within Multani Punjabi, thus contributing valuable insights to the field of linguistic evolution and variation within the Punjabi language. By elucidating the sociolinguistic dynamics influencing phonetic shifts and preservation of dialectal identity, this study will enrich our understanding of language change processes in diverse sociocultural contexts. Moreover, the findings may have implications for language revitalization efforts and policies aimed at preserving linguistic diversity within the Punjabi-speaking community.

Research Questions

(1) Which phonological changes are prevalent in the Multani Punjabi?
(2) Do the phonological alterations in Multani Punjabi exhibit any recognizable patterns, such as the merging or splitting of phonemes?
(3) What is the index of phonemic comparison between Multani Punjabi and standard Punjabi (Mahji)?
(3) What sociolinguistic variables and underlying causes are influencing the phonological changes in Multani Punjabi?
(4) What environmental elements affect Multani Punjabi word phonological changes, differentiating between conditioned and unconditioned changes?
In summary, there are notable phonological changes in the Punjabi Lexicon of Multani Punjabi, which may be ascribed to several linguistic and historical reasons. Various changes show how various dialects differ from one another and illustrate how intricate language change and evolution are.

Significance of the study

This work is important because it adds to our understanding of Multani Punjabi's changing linguistic environment, especially in light of the phonological alterations that have been noted. The research contributes important insights into language development, preservation, and identity within the Punjabi language family by recording these changes. The results could guide language planning and policy actions as well as attempts to preserve linguistic diversity.

Limitations of the study

Geographic Scope: Speakers of Multani Punjabi in Faridabad, Ballabhgarh, Palwal, and the surrounding areas are the primary focus of this study. The results might not accurately reflect phonological differences in Multani Punjabi in all areas of the world where the dialect is spoken.

Sample Size and Representativeness: The study's intentional sampling of speakers from particular regions may restrict the applicability of its conclusions to the larger community of Multani Punjabi speakers.

Time Restrictions: Enough time must be allotted for the gathering, processing, and interpretation of data in order to conduct a comprehensive study of phonological alterations. The study's scope and depth may be impacted by time restrictions.

Literature Review

The dynamic process of language evolution is shaped by a number of sociolinguistic, historical, and geographic variables. The purpose of this review of the research is to summarize the main conclusions from pertinent studies that provide insight into the phonological changes seen in Multani Punjabi, with an emphasis on tone loss and its consequences. Spoken in areas like Palwal, Ballabhgarh, and Faridabad, Multani Punjabi is distinguished from other Punjabi dialects by its distinct phonetic characteristics. Pitch contour variations are a major factor in differentiating between lexical meanings in traditional Punjabi dialects. (Rafi, 2010). Recent research, however, suggests that Multani Punjabi gradually loses tone, which begs the question of what forces this phonological change. Multani Punjabi phonological alterations are shaped in large part by language contact. Multani Punjabi's phonetic patterns and tonal variations have probably evolved as a result of historical interactions with nearby languages and cultures, such as Hindi and Haryanvi (Rafi, 2010). Understanding the intricate dynamics of Multani Punjabi phonological evolution can be gained by following the historical trajectory of language interaction. Understanding the observed phonological alterations requires an understanding of the sociolinguistic context of Multani Punjabi. Multani Punjabi speakers' language use and identity are influenced by urbanization, migration trends, and changing demography (Alizai, 2021). Sociolinguistic research emphasizes how societal changes affect phonological characteristics and how important it is to place phonological changes in the context of larger sociocultural processes. The lexicon and semantics of Multani Punjabi are significantly impacted by the lack of tone. Research indicates that tone leveling influences word meanings and subtleties, which may result in lexical variances and semantic ambiguity within the dialect (Rafi, 2010). Investigating the lexical effects of tone loss offers important new understandings of how Multani Punjabi phonological traits interact with linguistic semantics. Research comparing Multani Punjabi to Majhi and Dhani, two other Punjabi dialects, provides a more complex knowledge of the phonetic differences among the Punjabi language family. The vocabulary distinctions between Majhi and Dhani dialects are emphasized by Hasan (2021), underscoring the significance of dialectal variances in forming phonological traits and lexical richness within Punjabi dialects.

The recording of phonological alterations in Multani Punjabi has important ramifications for policy interventions and language preservation. Language planning activities that aim to preserve linguistic diversity and support the survival of endangered dialects such as Multani Punjabi can benefit from an understanding of the variables influencing phonological alterations. Protecting Multani Punjabi culture and identity requires policy interventions centered on language revitalization and community involvement.

These study frameworks make it easier to gather and analyze data, allowing scientists to investigate phonological changes and their implications in certain sociolinguistic contexts in a methodical manner.

Methodology

Nature of the Research

This comprehensive study adopts a mixed-methods approach to investigate phonological changes in Multani Punjabi, aiming to capture both the linguistic intricacies and sociolinguistic contexts of the dialect. The qualitative research methods employed enable a nuanced exploration of phonetic subtleties and cultural influences that shape language use within the Multani Punjabi-speaking community.

The Sample and Population

The study focuses on Multani Punjabi speakers residing in Faridabad, Ballabhgarh, Palwal, and surrounding areas, situated in the southern region of Punjab. A purposive sampling technique is employed to select 20 participants, evenly divided between Multani Punjabi speakers and those who speak standard Majhi Punjabi, which is considered a representative form of the Punjabi language. This sampling strategy ensures representation from various linguistic backgrounds within the Punjabi language family, allowing for a comprehensive analysis of phonological variations. The participants chosen were close relatives, friends, neighbours and family. This was to avoid the "observers' paradox" (Labov 1972) and elicit naturalistic speech.

Data collection methods include participant interviews, observations, and analysis of historical glossary information sourced from the "Glossary of the Multani Language in Comparison with Punjabi and Sindhi." Through a combination of conversational exchanges and archival data, the study aims to capture linguistic patterns and phonetic subtleties inherent in Multani Punjabi speech, with a particular focus on regional variants and unique linguistic traits.

Framework

The research framework emphasizes the identification and examination of Multani Punjabi's regional variants, focusing on language development and rule formation specific to the dialect's geographical context. By integrating qualitative data collection methods, fieldwork strategies, and an extensive literature review, the study aims to elucidate phonological changes and unique linguistic traits characteristic of Multani Punjabi.

For Phonemic Comparison, the Levenshtein algorithm will be used to determine the index, and further examine the data. Audio recordings of natural speech are transcribed and meticulously analyzed to identify noteworthy phonological changes, intonation patterns, and linguistic usage variations distinctive to Multani Punjabi. Comparative analysis is conducted by contrasting fieldwork findings and archival data with existing literature on the evolution of the Punjabi language, thereby contextualizing linguistic transformations within broader language studies.

Instruments and Data Collection

Voice samples are collected using phone recorders to capture a range of phonological elements from both Majhi and Multani Punjabi speakers. Participants engage in conversational tasks designed to elicit natural speech patterns, facilitating the examination of phonetic features and language variation.

Praat software, a powerful tool for phonetic analysis, is utilized for audio transcription and detailed phonetic examination. This software enables researchers to study Multani Punjabi pronunciation nuances, including intonation, stress patterns, and linguistic variations. The analysis of audio data using Praat facilitates the identification of phonological changes and linguistic innovations unique to the Multani Punjabi dialect.

Coding and Transcription

Phonetic transcription and thorough annotation of audio recordings are conducted using Praat software. Phonetic coding involves labeling phonological features, segmenting speech into phonetic units, and identifying rule formations and linguistic patterns in Multani Punjabi. Through systematic coding and transcription, the study aims to identify change patterns and rule

formations underlying phonological modifications in Multani Punjabi. This methodological approach integrates qualitative data analysis techniques with phonetic coding to advance our understanding of language variation and development within the Punjabi language family.

Comparative Analysis and Setting

The study conducts a comparative analysis of fieldwork findings, archival data, and existing literature on Punjabi language evolution to contextualize linguistic changes observed in Multani Punjabi. By contrasting linguistic data with historical records and contemporary scholarship, the research aims to deepen insights into the phonological dynamics of the Multani Punjabi dialect and its relation to broader language trends.

Ethical Considerations

Ethical considerations are paramount in this study, ensuring participant confidentiality, informed consent, and cultural sensitivity throughout the research process. The research adheres to ethical guidelines and protocols established by academic institutions and research organizations to protect the rights and well-being of study participants.

Data Analysis

Data analysis involves a multifaceted approach that integrates qualitative data interpretation with phonetic coding and rule formation analysis. The study employs thematic analysis to identify recurring patterns and themes in linguistic data, supplemented by detailed phonetic analysis using Praat software.

Limitations and Challenges

The study acknowledges potential limitations and challenges, including sample size constraints, language barriers, and logistical considerations in data collection. Strategies are implemented to address these challenges and ensure the rigor and validity of the research findings.

Significance and Contributions

This research contributes to the field of linguistics by providing a detailed examination of phonological changes in Multani Punjabi and their implications for language variation and development within the Punjabi language family. The study offers valuable insights into regional dialectal variations, sociolinguistic factors, and language contact phenomena, advancing our understanding of language diversity and evolution in multicultural societies.

In summary, this methodology outlines a rigorous research approach that integrates qualitative methods, linguistic analysis, and comparative studies to investigate phonological changes in Multani Punjabi. By leveraging advanced technologies and interdisciplinary frameworks, the study aims to generate meaningful insights into the dynamic nature of language variation and development within regional Punjabi dialects.

Data Collection and Analysis

Levensthein algorithm

The phonemic comparison of the Multani Punjabi and Mahji will be performed based on the Levensthein algorithm. By calculating the number of insertions and substitutions required to change one string into another, an algorithm can be used to quantify the difference between two strings. In 1965, Russian scientist Vladimir Levenshtein developed this technique. Greater

differences between strings are indicated by a higher Levenshtein distance. For instance, think about changing the word "house" to "mouse":
Replace "h" with "m" to make "house" become "mouse".
To finish the transformation, no more insertions or substitutions are needed.

The Levenshtein distance in this case between "house" and "mouse" is 1, indicating that only one substitution is required to change one word into the other. This algorithm offers a methodical approach to quantify the degree of variation between strings, making comparisons and analyses easier in a variety of information retrieval and computational linguistics applications.

Data Collected
Words that are common in speech of the individuals from both dialect and showed some differences in pronunciation have been selected from the data collected from listening to the audios of the participants:

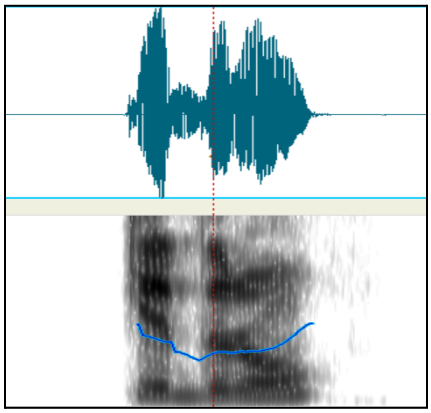
Mahji Words	**Mutlani Words**	**English Gloss**
Tussin	Tussan	You
Ki	Ke	What
Karda	Karende	Doing
Ain	E	Is
Kidron	Kitho	Where
Mera	Mainda	Mine
Nay	Naa	No
Saan	Saa	Breath
Chahida	Chaida	Desire
Karna	Karan	To do
Maran	Marenda	Seeking
Saran	Sarenda	To seek
Golan	Ghulenda	To burn
Balan	Balenda	Kindling
Kapan	Kaphijan	To cut
Dand	dand	Teeth
Din	Den	Day
Dena	Devan	To give

Table 1: *The words selected from the collected data from the audio*

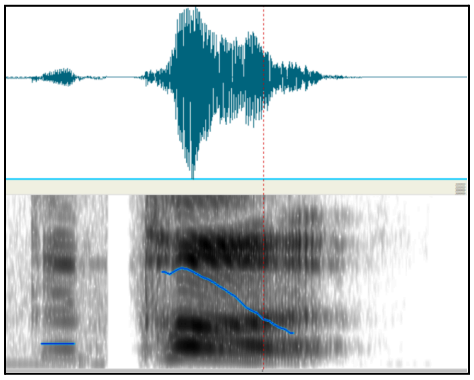
PRAAT software was used to transcribe the audios, and examine the pronunciation differences between the speech samples. For example, these are the frames of the word “Kidron”, Kithon”, “maran”, “marende”, “tussi” and “tussan”.

PRAAT transcriptions examples

(figure1 of speech sample)

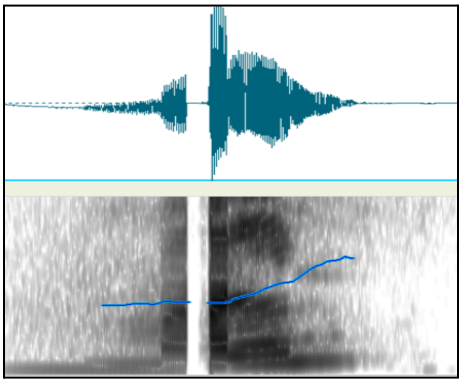


/k̄id̄rõ/

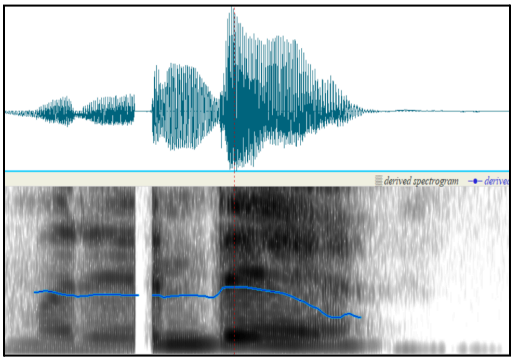


/ k̄īt̄ʰo/

(figure 2 of speech sample)

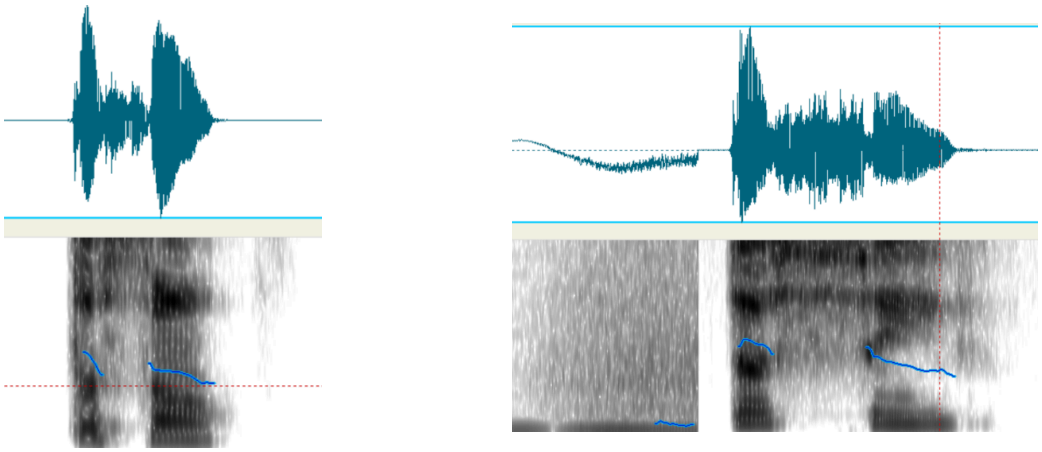


/ maȓan/



//mareñd̄e /

(figure 3 of speech sample)



Application of the Algorithm

The Levensthein algorithm's formula:

(figure 4 of the formula)

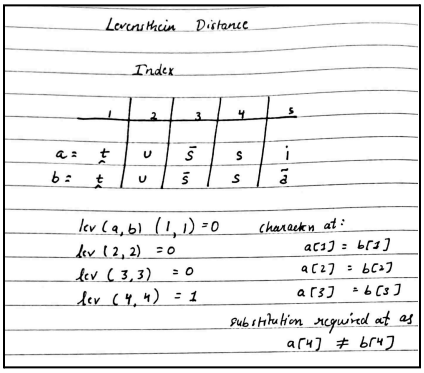
$$\operatorname{lev}(a,b) = \begin{cases} |a| & \text{if } |b| = 0, \\ |b| & \text{if } |a| = 0, \\ \operatorname{lev}\left(\operatorname{tail}(a), \operatorname{tail}(b)\right) & \text{if } \operatorname{head}(a) = \operatorname{head}(b), \\ 1 + \min \begin{cases} \operatorname{lev}\left(\operatorname{tail}(a), b\right) \\ \operatorname{lev}\left(a, \operatorname{tail}(b)\right) \\ \operatorname{lev}\left(\operatorname{tail}(a), \operatorname{tail}(b)\right) \end{cases} & \text{otherwise} \end{cases}$$

Implementing the algorithm, the phonemic comparison has been made between the two dialects. An example to understand how the algorithm is applied:

The word of Mahji "tussin" and the word of Multani "tussan", the Levensthein distance is 1 since there is substitution of / ĩ/ "in" with /ã/ "an".

To understand, how the index is calculated, here is the calculation based on the formula of the word "tussin" and "tussan"

(figure 5 the application of formula)


Levensthein Distance
Index
1 2 3 4 5
a = t̃ u s̄ s i
b = t̃ u s̄ s ã
lev (a, b) (1, 1) = 0 characters at:
lev (2, 2) = 0 a[1] = b[1]
lev (3, 3) = 0 a[2] = b[2]
lev (4, 4) = 1 a[3] = b[3]
substitution required at as
a[4] ≠ b[4]

The following analysis focuses on the phonetic sounds (consonants) extracted from the provided data comparing Multani and Majhi dialects of Punjabi, using the Levenshtein distance algorithm to quantify similarities and differences.

Similar Consonant Sounds (Levenshtein Distance 0)
(table 1.2)

Phoneme in Punjabi	Phoneme in Multani	Levensthein distance
/t̄/	/t̄/	0
/m̄/	/m̄/	0
/k̄/	/k̄/	0
/s̄/	/s̄/	0
/t͡ʃ̄/	/t͡ʃ̄/	0
/b̄/	/b̄/	0

Consonant Sounds with One Character Difference (Levenshtein Distance 1)
(table 1.3)

Phoneme in Punjabi	Phoneme in Multani	Levensthein distance
/d̄/	/d/	1
/d/	/tʰ/	1
/r̥/	/r/	1
/p/	/pʰ/	1
/n/	/w/	1

Vowels Sounds with One Character Difference (Levenshtein Distance 0)
(table 1.4)

Phoneme in Punjabi	Phoneme in Multani	Levensthein distance
/a/	/a/	0
/o/	/o/	0

Vowels Sounds with One Character Difference (Levenshtein Distance 1)
(table 1.5)

Phoneme in Punjabi	Phoneme in Multani	Levensthein distance
/æ/	/ə/	1
/e/	/a/	1
/ã/	/a/	1
/i/	/a/	1
/i/	/ə/	1
/e/	/ə/	1

Calculation

Similar Consonant Sounds (Levenshtein Distance 0):

- Total sounds with zero Levenshtein distance: 6

Consonant Sounds with One Character Difference (Levenshtein Distance 1):

- Total sounds with Levenshtein distance of 1: 5

Vowel Sounds Analysis:

Vowel Sounds with Zero Character Difference (Levenshtein Distance 0):

- Total vowel sounds with zero Levenshtein distance: 2

Vowel Sounds with One Character Difference (Levenshtein Distance 1):

- Total vowel sounds with Levenshtein distance of 1: 6

Calculation of Similarity and Difference Indices:

1. Similarity Index:
    - Total Similar Sounds (Levenshtein Distance 0): 8 (6 consonants + 2 vowels)
    - Total Possible Sounds (consonants + vowels): 13 (6 consonants + 7 vowels)
    - Similarity Index = (Total Similar Sounds / Total Possible Sounds) * 100
    - Similarity Index = (8 / 13) * 100 ≈ 61.54%
2. Difference Index:
    - Total Different Sounds (Levenshtein Distance 1): 11 (5 consonants + 6 vowels)
    - Difference Index = (Total Different Sounds / Total Possible Sounds) * 100
    - Difference Index = (11 / 13) * 100 ≈ 84.62%

Phonemic Comparison Ratio:

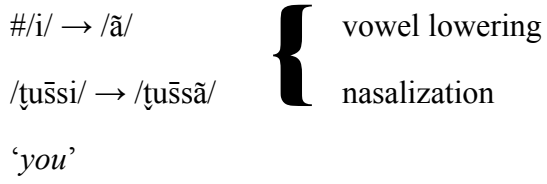
- Phonemic Similarity Ratio: 61.54%
- Phonemic Difference Ratio: 84.62%

These indices and ratios provide a quantitative measure of the linguistic variation between Majhi and Multani dialects of Punjabi based on the Levenshtein distances of consonant and vowel sounds. The similarity index indicates the proportion of identical sounds, while the difference index highlights the proportion of sounds with slight variations between the dialects.

Phonological Rules

The upcoming section will explain the phonological rules governing word formation in Multani Punjabi. It will explore how phonetic shifts and rule patterns shape the dialect's linguistic structure. The analysis will focus on systematic transformations like assimilation, deletion, and modification in Multani Punjabi, shedding light on language variation and change. The phonological changes that took place in the Multani Punjabi, follow certain rule formation patterns. These alternations can be divided into a few phonological processes' categories such as vowel raising, vowel lowering, assimilation, addition, deletion, and substitution. Based on the changes recorded from the data collection of this study, following are the rule formations.

Vowel Lowering, Nasalization, Phonemic length addition



/ʈuɞ̄si/ ('you') undergoes two changes to become /ʈuɞ̄sã/:

1. The high front vowel /i/ in /ʈuɞ̄si/ is lowered to /a/ in /ʈuɞ̄sã/.
2. The vowel /a/ is nasalized, resulting in /ã/.

This analysis shows how specific phonetic rules apply to modify the structure of words within a language. The processes of vowel lowering and nasalization work together to change the vowel sound in the word 'you' from /ʈuɞ̄si/ to /ʈuɞ̄sã/.

#/i/ → /ə/
/ki/ → /ə/ } vowel lowering

‘*what*’

/ki/ (‘what’) undergoes a vowel change to become /ke/:
The high front vowel /i/ in /ki/ is lowered to the mid-front vowel /e/ in /ke/.The process of vowel lowering changes the vowel sound in the word ‘what’ from /ki/ to /ke/.

#/e/ → /a:/ } vowel lowering
/n̄e/ → /n̄a:/ } Phonemic length addition

‘*no*’

The process of vowel lowering changes the vowel sound in the word ‘what’ from /ki/ to /kə/,
/nə/ (‘no’) undergoes two changes to become /na:/:

1. The mid-front vowel /e/ in /ne/ is lowered to the low-back vowel /a/ in /na/.
2. The vowel /a/ is lengthened, resulting in /a:/ in /na:/.

The processes of vowel lowering and phonemic length addition work together to change the vowel sound in the word ‘no’ from /n̄e/ to /n̄a:/

Vowel Raising, Addition, Assimilation, Declusterisation

#/a/ → #/ə/
/k̄ərda:/ → /k̄ərd:ende/ } vowel raising, addition, assimilation, declusterisation

‘*doing*’

/k̄ərda:/ (‘doing’) undergoes four changes to become /k̄ərdende/:

1. The low central vowel /a/ in /k̄ərda:/ is raised to the mid-front vowel /ə/ in /k̄ərdende/.
2. The segment /nd/ is added to the original word.
3. The newly added segments assimilate to the existing vowel sounds /e/, creating a cohesive phonological structure.

The processes of vowel raising, addition, and assimilation work together to change the vowel sound and overall structure of the word ‘doing’ from /k̄ərda:/ to /k̄ərdende/

Denasalization, Vowel Lowering

#/æ̃/ → /e/ } Vowel raising, Denasalization
[“Ain” and “E”]

‘*is*’

‘æ̃’ (‘is’) undergoes two changes to become ‘e’:

1. The low front nasalized vowel /æ̃/ is raised to the mid-front vowel /e/.
2. The nasalization is removed, resulting in the non-nasal vowel /e/.

The processes of vowel raising and denasalization work together to change the vowel sound in the word ‘is’ from ‘æ̃’ to ‘e’

#/a̰:/ → /a:/ denasalization

/sa̰:/ → /sa:/

‘*breath*’

/sa̰:/ (‘breath’) undergoes one change to become /sa:/:
The nasalized vowel /a̰:/ in /sa̰:/ is denasalized, resulting in the non-nasal vowel /a:/ in /sa:/.

Deletion, Aspiration

#/k̄idrõ/ → /k̄iṯʰõ/ Deletion, Dental plosive conversion, Aspiration

‘*Where*’

/k̄idrõ/ (‘where’) undergoes three changes to become /k̄iṯʰõ/:

1. The sound /r/ is deleted from /kidrõ/, resulting in /k̄idõ/.
2. The plosive /d/ in /k̄idõ/ is converted to the dental plosive /ṯ/, resulting in /k̄itõ/.
3. The dental plosive /ṯ/ is aspirated, resulting in /k̄iṯʰõ/.

The processes of deletion, dental plosive conversion, and aspiration work together to change the sound in the word ‘where’ from /k̄idrõ/ to /k̄iṯʰõ/

#/t͡ʃahiḍa/ → / t͡ʃaiḍa/ Deletion

‘*want*’

/t͡ʃahiḍa/ (‘want’) undergoes one change to become /t͡ʃaiḍa/:

1. The sound /h/ is deleted from /t͡ʃahiḍa/, resulting in /t͡ʃaiḍa/.

The process of deletion changes the word ‘want’ from /t͡ʃahiḍa/ to /t͡ʃaiḍa/

Metathesis

#/k̄arn̄a/ → /k̄ar̄an/ Metathesis

The sounds /n/ and /a/ are transposed. In /k̄arn̄a/, the sequence /na/ is changed to /an/, resulting in /k̄ar̄an/.
This analysis demonstrates how a specific phonetic rule applies to modify the structure of words within a language. The process of metathesis changes the word ‘to do’ from /k̄arn̄a/ to /k̄ar̄an/

#/deñ̄a/ → /devañ̄/ Metathesis, Addition

1. “n̄a” changed to “añ̄”, where the transposition of two phonemes took place.
2. There is the addition of /v/ in the word median position, resulting in the change, /deñ̄a/ to /devañ̄/.

Addition (Conditioned Change)

The pattern that was noticed in the phonological alteration was that, in case of intervocalic words, the addition of the segment "da" gets added at the word final position. The pattern is noticed in the following words that were collected during the speech analysis. Along with the addition of "da", there are other patterns also noticed in the word change.

The addition of "da" is applicable to the following listed formations.

#/mərən̄/ → /mərenda { Dissimilation

/ə/ changed due to the neighbouring influence into the phoneme /ə/, the phoneme got dissimilar to the existing vowel / ə/ at the word median position.

Same formation happened with the word:

#/bəl̄ən̄/ → /bəl̄enna/ { Vowel lowering

#/gol̄ən̄/ → /gʰul̄enda/ { Aspiration

1. /g/ → /gʰ/
2. /o/ → /u/
3. /ə/ → /e/

{ Aspiration process took place, Vowel lowering process

#/kəp̄ən̄/ → /kəp̄id͡ʒan̄/ { Dissimilation, Addition

1. The phoneme /ə/ changed to /i/, under the influence of the neighbouring phoneme, dissimilarity was invoked.
2. The addition of /d͡ʒ/ took place in the word median position.
3. The phoneme /ə/ changed to /a/, under the influence of the neighbouring phoneme, and dissimilarity was invoked.

Analysis of the phonological changes

Vowel Lowering, Nasalization, Phonemic Length Addition

1. Vowel Lowering and Nasalization:
   - Change: /i/ → /ã/
   - Example: /ʈus̄si/ → /ʈus̄sã/ ('you')
   - Explanation: The high front vowel /i/ is lowered to a low central vowel /a/, and then nasalized to /ã/. This change can be attributed to a natural tendency in phonetic evolution where high vowels are sometimes lowered for ease of articulation, and nasalization can occur due to the influence of nasal consonants or phonetic context.
2. Vowel Lowering:
   - Change: /i/ → /e/
   - Example: /ki/ → /ke/ ('what')
   - Explanation: The high front vowel /i/ is lowered to the mid-front vowel /e/. Vowel lowering might be driven by the need to create a more open vowel sound, which

is often easier to articulate and can be influenced by surrounding phonetic contexts.

2. Vowel Lowering and Phonemic Length Addition:
    - Change: /ə/ → /a:/
    - Example: /n̄ə/ → /n̄a:/ (‘no’)
    - Explanation: The mid-front vowel /ə/ is lowered to a low-back vowel /a/ and then lengthened to /a:/. Lengthening can occur to maintain rhythmic balance in speech or to distinguish between words (phonemic contrast).

Vowel Raising, Addition, Assimilation, Declusterization

1. Vowel Raising, Addition, Assimilation:
    - Change: /a/ → /e/ and Addition of /nd/
    - Example: /k̄ərda:/ → /k̄ərdende/ (‘doing’)
    - Explanation: The low central vowel /a/ is raised to the mid-front vowel /e/. Additionally, the segment /nd/ is added and assimilates to the existing vowel sounds, creating a phonologically cohesive structure. These changes can be attributed to the tendencies for creating more varied vowel sounds (raising) and enhancing syllable structure through addition and assimilation.

Denasalization, Vowel Raising

1. Denasalization and Vowel Raising:
    - Change: /æ̃/ → /e/
    - Example: ‘æ̃’ → ‘e’ (‘is’)
    - Explanation: The nasalized vowel /æ̃/ is raised to the mid-front vowel /e/ and denasalized. Denasalization removes the nasal quality, possibly due to a loss of a nasal environment or simplification in speech.
2. Denasalization:
    - Change: /a:̃/ → /a:/
    - Example: /sa:̃/ → /sa:/ (‘breath’)
    - Explanation: The nasalized vowel /a:̃/ loses its nasal quality, resulting in /a:/. This change often occurs in the absence of nasal consonants that initially caused the nasalization.

Deletion, Aspiration

1. Deletion, Dental Plosive Conversion, Aspiration:
    - Change: /kidrõ/ → /kiṯʰõ/
    - Example: /kidrõ/ → /kiṯʰõ/ (‘where’)
    - Explanation: The sound /r/ is deleted, the plosive /d/ is converted to a dental plosive /ṯ/, and then aspirated. Deletion simplifies the word structure, dental conversion might reflect dialectal influences, and aspiration can enhance the perceptual distinctiveness of the consonant.

Metathesis

1. Metathesis:
    - Change: /kaɾ̄na/ → /kaɾ̄an/
    - Example: /kaɾ̄na/ → /kaɾ̄an/ (‘to do’)

- Explanation: The sounds /n/ and /a/ are transposed. Metathesis often occurs due to ease of articulation or perceptual factors, where speakers naturally transpose sounds to produce a more fluent speech pattern.

Metathesis, Addition

1. Metathesis and Addition:
    - Change: /deña/ → /devañ/
    - Example: /deña/ → /devañ/ (‘give’)
    - Explanation: The sequence /ña/ is transposed to /añ/, and /v/ is added. These changes could reflect analogical processes where a pattern of adding /v/ became regularized, or due to phonological environments that favoured such transitions.

Addition (Conditioned Change)

1. Addition and Dissimilation:
    - Change: /mərəñ/ → /məreñda/
    - Example: /mərəñ/ → /məreñda/ (‘to die’)
    - Explanation: The addition of /da/ occurs in intervocalic words, with dissimilation where /ə/ changes to /e/ due to neighboring influences. These changes might be driven by analogical processes and phonological conditioning where a specific suffix (/da/) is added to conform to a morphological pattern.
2. Addition, Aspiration, Vowel Lowering:
    - Change: /gol̄əñ/ → /gʰul̄eñda/
    - Example: /gol̄əñ/ → /gʰul̄eñda/ (‘to swallow’)
    - Explanation: The changes include aspiration (/g/ to /gʰ/), vowel lowering (/o/ to /u/), and /ə/ changing to /e/. These alterations may be attributed to analogical influences, assimilation, and conditioned phonological environments.
3. Addition and Dissimilation:
    - Change: /kəp̄əñ/ → /kəp̄id͡ʒañ/
    - Example: /kəp̄əñ/ → /kəp̄id͡ʒañ/ (‘to jump’)
    - Explanation: The changes involve /ə/ to /i/ and /a/ (dissimilation), and the addition of /d͡ʒ/. This might reflect a pattern where complex syllable structures favor the addition of new segments to maintain a phonological balance.

Sociolinguistics factor behind the phonological changes

Historical Sound Shifts' Impact

When compared to Majhi Punjabi, Multani Punjabi's phonological traits have been significantly shaped by historical sound shifts. These shifts show how pronunciation patterns have changed systematically over time, frequently as a result of social and linguistic influences.

Vowel reduction or elision is a common phenomenon observed in many languages, including Punjabi, and is a key component of historical sound shifts. The shift from "karda" (to do) to "karende" in Multani Punjabi is an example of this phenomenon. When conjugated in a specific grammatical context, the final vowel "a" in "karda" is reduced or elided, creating the form "karende". This kind of phonological alteration is related to historical linguistic processes in which vowels are changed to fit into grammatical structures or linguistic patterns.

Furthermore, assimilation or pronunciation modifications of consonants may be part of historical sound shifts. For example, the change from "Din" (day) (Majhi) to "den" (Multani) indicates a change in consonant articulation over time. Punjabi dialects frequently experience

assimilation or simplification of consonant clusters over time, resulting in regional differences in pronunciation.

Furthermore, interaction with other languages or dialects can impact historical shifts in sound, leading to phonological innovations. Because Multani Punjabi is spoken in an area with many different linguistic influences, phonetic changes may have occurred as a result of contact with nearby languages or older dialectal forms. Multani Punjabi has a unique phonetic profile due in part to the adoption of new phonological traits or the modification of preexisting ones brought about by this interaction. Studies in Punjabi linguistics, like those by Gupta (2011), offer important new perspectives on the historical shifts in sound that have been noted in various Punjabi dialects. These studies illustrate the dynamic nature of language change and adaptation within the Punjabi language family by charting the evolution of phonological patterns over time.

Regional and Dialectal Variation

The phonological characteristics and linguistic diversity between Multani Punjabi and Majhi Punjabi are largely shaped by regional dialectal variation. The pronounced variations in phonological traits among these dialects are indicative of the impact of geographical dynamics and historical regional factors. Words like "Kidron" (where) (Majhi) become "Kitho" (Multani), demonstrating the influence of regional dialectal variation on phonological characteristics. Compared to the standard Majhi Punjabi spoken in central Punjab, Multani Punjabi, which is spoken in the southern part of the state, has distinct phonetic characteristics. The linguistic divergence that has occurred over time can be ascribed to historical migrations, settlement patterns, and geographic isolation.

Take the switch from "Tussin, (you) assin" (I/us) (Majhi Punjabi) to "Tussan, assan" (Multani Punjabi) as an example. The change from "i" to "a" in the pronouns "Tussin" and "assin" is indicative of a shared phonological pattern impacted by nearby dialects. This shift may be explained by phonetic assimilation, in which Multani Punjabi takes on vowel sounds more typical of nearby dialects like Saraiki or Sindhi, which frequently use the sound "a" rather than "i" in specific situations. Hussain's (2011) study of Saraiki linguistic characteristics may shed light on these phonological parallels.

The change from "Karna" (to do) (Majhi Punjabi) to "Karan" (Multani Punjabi) is another example of how nearby dialectal characteristics influence phonetic adaptation. In Multani Punjabi, the final vowel in the verb "Karna" changes from "a" to "an" at the end. This change in pronunciation may be due to phonological influences from neighboring dialects like Haryanvi.

Furthermore, lexical assimilation and phonological adaptation are highlighted by the evolution of words like "Golan" (to kindle) (Majhi) to "Gulenda" (Multani). Words in Multani Punjabi may have changed consonant clusters and vowel sounds due to influences from nearby dialects, such as Sindhi, which also shows phonetic shifts.

Language Contact

The observed word differences between other Punjabi dialects and Multani Punjabi can be explained by the phenomenon of language contact, which is impacted by interactions with nearby languages like Dobari, Hindi, Pahari, and Potohari. In the past, these interactions have helped Punjabi-speaking areas adapt their language and undergo phonological changes.

(figure 6 of types of language contact processes)

Type	Process	Description	Speakers/bilingualism	Language attitudes
A	Phonological adaptation of word-forms	Replicated word-forms are adjusted to match the sound patterns of the recipient language	Semi-bilinguals or monolinguals	Strong loyalty towards, and stability of the recipient language; superficial contact
B	Borrowing of phonological features along with word-forms	Borrowed and inserted word-forms maintain (fully or partly) the original sound patterns of the donor language ('authentication')	Fairly widespread bilingualism	Flexibility in the use of the recipient Language, prestigious bilingualism
C	Convergence of systems during second-language acquisition	Word-forms of the target language are systematically adjusted to match the sound patterns of the native language	Emerging bilingualism; stable minority bilingualism; emergence of ethnolect or language shift	Strong group identity coupled with a need (pressure) to acquire the target language
D	Convergence of systems in stable, intensive bilingualism	Sound patterns of the native language are adjusted to match those of the second language	Intensive and widespread bilingualism	Second language is 'prestige' language

Text taken from Phonological changes in the Hindi lexicon: a case of Meghalaya Hindi by Maansi Sharma

Phonological Assimilation and Language Contact: The change of words such as "Tussin, assin" (Majhi Punjabi) to "Tussan, assan" (Multani Punjabi) is an example of how language contact influences phonological assimilation. Multani Punjabi may take on vowel sounds that are more typical of nearby languages like Hindi, where "a" is frequently used in place of "i" in some pronouns. Scholars have studied Punjabi and Hindi language contact and have shown how cross-linguistic influences affect phonological features.

Word Evolution from "Golan" (to burn) (Majhi) to "Gulenda" (Multani): This shows how nearby dialects like Pahari and Potohari have influenced both lexical borrowing and phonological adaptation. Because these dialects are geographically close to Multani Punjabi-speaking areas, phonetic patterns and sounds have been incorporated into Multani vocabulary.

Loans and Borrowings

The phonological structure of Multani Punjabi has been significantly impacted by the intermingling of loanwords from other languages. Loanwords are lexical terms that are adopted from one language into another, frequently to fill lexical gaps or for reasons of linguistic prestige, as noted by Trudgill (1992) and Crystal (2008). Owing to its location in an area with a wide range of linguistic influences, Multani Punjabi has undergone phonological changes as a result of absorbing phonetic elements from languages such as Hindi, English, Arabic and even Urdu.

For example, words borrowed from Haryanvi or Hindi may change the phonetic patterns of Multani Punjabi or introduce new phonemes. As a result of the adoption of these loanwords, phonemic distribution and pronunciation may alter, reflecting the dynamics of ongoing language contact in the area (Rahman, 2006). "Paradise" or "heaven" is Jannat (جنت), pronounced as /d͡ʒənnəʈ̪/ in Multani Punjabi, with the first /d͡ʒ/ sound being pronounced with Arabic influence. Namaz (نماز) is an Arabic word that means "prayer" in reference to Islamic prayer. In Multani Punjabi, the word is pronounced as /nəmā:z/. Hukam (حاما) has the meaning "order" or "command" and is pronounced as hukam in Multani Punjabi, with the initial "h" sound being

influenced by Urdu. Shayari (شاعرó) is a word that means "poetry" and is pronounced in Multani Punjabi as /ʃāʿirī/, which is how the word is pronounced in Urdu. School, pronounced as /skūl/ in Multani Punjabi, is a word meaning "school" that reflects the phonological simplification of English loanwords. This word illustrates how Punjabi phonology has adapted English loanwords.

Media Consumption

The conversion of pronouns like "Tussin, assin" (Majhi Punjabi) to "Tussan, assan" (Multani Punjabi) is one prominent example of phonological change brought about by media in Multani Punjabi. Pronoun usage may change from "i" to "a" as a result of exposure to local media broadcasts or entertainment shows that feature speakers of a particular language. Regular media consumption by Multani Punjabi speakers may cause them to subconsciously absorb the phonetic patterns of celebrities or actors in the media, changing their pronunciation over time.

Similar to how vocabulary and pronunciation can change due to media exposure, this can also be observed in variants such as "Karna" (Majhi Punjabi) becoming "Karan" (Multani Punjabi). Multani Punjabi speakers' language usage may be influenced by the adoption of specific phonetic patterns, such as the pronunciation of final vowels, which correspond with popular media representations.

Language accommodation and the prestige attached to particular speech patterns that are promoted by media outlets are the causes of media-induced phonological changes. Certain linguistic elements heard in media content may be interpreted by Multani Punjabi speakers as prestigious or modern, prompting the adoption of these elements into everyday speech.

Studies on the relationship between language and media, like those by Gershon, I. & Manning, Paul. (2014) investigate the ways in which media exposure affects language change and variation in multilingual settings. Gupta's research demonstrates how media consumption influences the linguistic patterns and phonetic modifications that speakers of regional dialects, such as Multani Punjabi, make.

The well-known Multani Punjabi film "Choorian," helmed by Syed Noor, is one example of a work that has improved linguistic representation. Audiences are exposed to the phonetic features of Multani Punjabi dialect through this film's soundtrack, which includes songs and dialogue performed in the dialect. The popularity of the movie has affected viewers' phonological preferences and language usage in addition to spreading awareness of Multani culture. Similar phonological characteristics and pronunciation patterns can be heard in Multani Punjabi songs particularly in those sung by regional performers like Arif Lohar and Attaullah Khan Esakhelvi. These songs serve as cultural artifacts that preserve and promote Multani Punjabi linguistic heritage because they frequently employ regional vocabulary and phonetic patterns.

Discussion

In this section, the answers to projected research questions will be answered.

(1) Which phonological changes are prevalent in the Multani Punjabi?

Phonological changes in Multani Punjabi reflect the dynamic nature of spoken language as it evolves through social and environmental influences. Prevalent changes include vowel lowering, where high vowels like /i/ are often lowered to mid or low vowels such as /a/ and /ə/. Nasalization and denasalization indicate shifts between nasal and non-nasal vowel forms. Other common changes involve phonemic length addition, aspiration of consonants, deletion of sounds for simplification, addition of segments in intervocalic positions, and metathesis, where sounds

within words are transposed. These changes highlight the fluidity and adaptability of Multani Punjabi's phonological system. The prevalent phonological changes in Multani Punjabi include vowel lowering, nasalization, denasalization, phonemic length addition, aspiration, deletion, addition, and metathesis.

The prevalent phonological changes in Multani Punjabi include:

- Vowel Lowering: High front vowels like /i/ are lowered to mid or low vowels such as /a/ and /ə/ (e.g., /ki/ → /ke/).
- Nasalization: Vowels undergo nasalization in specific phonetic contexts (e.g., /ʈus̄si/ → /ʈus̄sã/).
- Denasalization: Nasal vowels become non-nasal (e.g., /sa̰:/ → /sa:/).
- Phonemic Length Addition: Vowels are lengthened to distinguish meaning or for rhythmic balance (e.g., /n̄e/ → /n̄a:/).
- Aspiration: Consonants become aspirated, particularly dental plosives (e.g., /d/ → /tʰ/).
- Deletion: Certain sounds are deleted to simplify word structure (e.g., /r/ in /k̄idrõ/ → /k̄iʈʰõ/).
- Addition: New segments are added, often to intervocalic positions (e.g., /mərən̄/ → /mərenda/).
- Metathesis: Transposition of sounds within a word (e.g., /k̄arn̄a/ → /k̄aran̄/).

(2) Do the phonological alterations in Multani Punjabi exhibit any recognizable patterns, such as the merging or splitting of phonemes?

Understanding phonological patterns in Multani Punjabi involves identifying systematic alterations in sound structure. These changes often follow recognizable patterns such as merging phonemes, which simplifies the phonetic inventory, and splitting phonemes, which creates new distinctions. Assimilation, where phonemes change to become more like neighboring sounds, is another significant pattern. Conditioned changes occur in specific phonetic environments, demonstrating the influence of surrounding sounds on phoneme alteration. These patterns provide insights into the structural evolution of the dialect and its adaptation to linguistic and social contexts.

The phonological alterations in Multani Punjabi exhibit recognizable patterns, including merging and splitting of phonemes, assimilation, and conditioned changes.

The phonological alterations in Multani Punjabi exhibit recognizable patterns, such as:

- Merging of Phonemes: Some phonemes are merged, simplifying the phonetic inventory. For instance, the distinction between nasalized and non-nasalized vowels can be reduced.

Original: /ã/ (nasalized vowel) and /a/ (non-nasalized vowel)

Merged: /a/

Explanation: The nasalized and non-nasalized versions of the vowel /a/ merge into a single phoneme, simplifying the phonetic inventory. This can happen due to a lack of phonemic distinction needed in everyday speech.

- Splitting of Phonemes: In some contexts, phonemes are split, creating new distinctions, particularly with vowel length or aspiration.

Original: /a/

Split: /a/ (short vowel) and /a:/ (long vowel)

Explanation: The original short vowel /a/ splits into two distinct phonemes, /a/ and /a:/, based on vowel length. This distinction helps in differentiating meanings of words in certain contexts.

- Assimilation: Phonemes assimilate to neighboring sounds to maintain phonological cohesion (e.g., /nd/ addition in /k̄ərdəndə/).

Original: /k̄ərda:/ (‘doing’)
Assimilated: /k̄ərdende/
Explanation: The addition of /nd/ results in assimilation to the existing vowel sounds, creating a more cohesive phonological structure. This process helps maintain phonological harmony within the word.

- Conditioned Changes: Specific environments condition certain changes, such as vowel lowering in the presence of nasal consonants.

Addition and Dissimilation
Example:
Original: /mərən̄/ (‘to die’)
Changed: /mərenn̄da/
Explanation: The addition of /da/ occurs intervocalically with dissimilation where /ə/ changes due to neighboring influences. This is driven by analogical processes and phonological conditioning to conform to morphological patterns.

(3) What is the index of phonemic comparison between Multani Punjabi and standard Punjabi (Mahji)?
Quantifying phonemic similarities and differences between Multani and Majhi dialects helps in understanding their linguistic relationship. Using the Levenshtein distance algorithm, the similarity index for phonemes shows that approximately 61.54% of the sounds are identical, while the difference index indicates that 84.62% of the sounds differ by one character. These indices reveal a moderate level of phonemic similarity, highlighting the distinct phonological characteristics of each dialect. This comparative analysis is crucial for linguists studying dialectal variation and language change.
The phonemic comparison between Multani Punjabi and standard Punjabi (Majhi) shows a similarity index of 61.54% and a difference index of 84.62%.

The phonemic comparison between Multani Punjabi and standard Punjabi (Majhi) is quantified using the Levenshtein distance algorithm:

- Similarity Index: 61.54% (based on 8 out of 13 sounds being identical).
- Difference Index: 84.62% (based on 11 out of 13 sounds differing by one character).

These indices indicate a moderate level of phonemic similarity with substantial phonological variation.

(4) What sociolinguistic variables and underlying causes are influencing the phonological changes in Multani Punjabi?
Sociolinguistic variables play a crucial role in shaping phonological changes in Multani Punjabi. Factors such as language contact, geographic isolation, social stratification, and cultural identity drive these changes. Interaction with neighboring languages and dialects leads to borrowing and assimilation of phonological features. Geographic isolation can result in unique phonological developments within the dialect. Social class, education, and occupation influence speech patterns, while a strong cultural identity promotes the retention of distinct phonological traits. These variables highlight the complex interplay between social factors and language evolution. Sociolinguistic variables influencing phonological changes in Multani Punjabi include language contact, geographic isolation, social stratification, and cultural identity.

Sociolinguistic variables and underlying causes influencing phonological changes in Multani Punjabi include:

- Language Contact: Interaction with neighboring languages and dialects influences phonological changes through borrowing and assimilation.
- Geographic Isolation: Regional isolation can lead to dialect-specific phonological developments.
- Social Stratification: Differences in social class, education, and occupation might influence speech patterns and phonological changes.
- Cultural Identity: A strong regional identity can promote the maintenance of distinct phonological features as a marker of cultural heritage.

(5) What environmental elements affect Multani Punjabi word phonological changes, differentiating between conditioned and unconditioned changes?

Environmental elements impacting phonological changes in Multani Punjabi can be divided into conditioned and unconditioned changes. Conditioned changes occur in specific phonetic environments, such as intervocalic positions leading to segment addition or vowel lowering influenced by neighboring sounds. Unconditioned changes, like simplification through deletion or transposition of sounds (metathesis), occur independently of surrounding phonetic context. Analyzing these changes helps linguists understand how phonetic context and structural ease drive phonological evolution, revealing the adaptability of the dialect.

Environmental elements affecting phonological changes in Multani Punjabi include conditioned changes, influenced by specific phonetic environments, and unconditioned changes, which occur independently of surrounding contexts.

Environmental elements affecting phonological changes in Multani Punjabi can be classified into conditioned and unconditioned changes:

Conditioned Changes:

- Intervocalic Environments: Addition of segments like "da" in word-final positions when surrounded by vowels (e.g., /mərən̄/ → /mərenn̄da/).
- Influence of Neighboring Phonemes: Changes like vowel lowering or dissimilation due to the influence of adjacent sounds (e.g., /ə/ to /ə/ in /kəp̄ən̄/ → /kəp̄id͡ʒan̄/).

Unconditioned Changes:

- Simplification Processes: Deletion of sounds to simplify word structure (e.g., /r/ deletion in /kidrõ/ → /kiʈ̪ʰõ/).
- Metathesis: Transposition of sounds that occurs independently of specific phonetic contexts (e.g., /k̄ar̄na/ → /k̄ar̄an/).

These changes are driven by both phonological ease and social influences, reflecting the dynamic nature of language evolution in Multani Punjabi.

Conclusion

The phonological analysis of Multani Punjabi, compared to standard Punjabi (Majhi), reveals significant insights into the linguistic dynamics at play. This study, grounded in the comparison of consonant and vowel sounds using the Levenshtein distance algorithm, highlights the prevalent phonological changes, recognizable patterns, phonemic comparisons, and the sociolinguistic and environmental factors influencing these changes.

Multani Punjabi exhibits several notable phonological changes. Vowel lowering is a prominent trend, where high vowels such as /i/ are lowered to mid or low vowels like /ə/ and /a/,

as seen in /ki/ becoming /kə/ ('what'). Nasalization and denasalization also play significant roles, transforming vowels to and from nasalized forms, respectively. For instance, /ʈus̄si/ changes to /ʈus̄sã/ ('you'). Additionally, phonemic length addition, where vowels are lengthened (e.g., /n̄ə/ to /n̄a:/ for 'no'), enhances rhythmic balance and phonemic distinction.

The phonological alterations in Multani Punjabi follow discernible patterns such as merging and splitting of phonemes. Merging simplifies the phonetic inventory by combining distinct phonemes, while splitting creates new distinctions, adding complexity. Assimilation, where phonemes change to become more like neighboring sounds (e.g., /k̄ərda:/ to /k̄ərdende/), creates cohesive phonological structures. These patterns underscore the structured yet adaptable nature of phonological evolution in the dialect.

The phonemic comparison between Multani and Majhi dialects, measured through the Levenshtein distance, reveals a similarity index of approximately 61.54%. This indicates that a significant portion of sounds are identical between the two dialects. However, the difference index of 84.62% highlights the extent of phonemic variation, with many sounds differing by one character. These indices provide a quantitative measure of the linguistic variation, illustrating both shared features and distinct characteristics.

Sociolinguistic variables significantly influence phonological changes in Multani Punjabi. Language contact introduces new phonological features, while geographic isolation fosters unique traits. Social stratification and cultural identity also shape speech patterns, reflecting the community's heritage and social dynamics. Environmental factors further drive phonological changes. Conditioned changes occur in specific phonetic contexts, influenced by neighbouring sounds, while unconditioned changes, such as deletion and metathesis, occur independently of phonetic context.

In conclusion, the phonological analysis of Multani Punjabi demonstrates a dynamic interplay of phonetic evolution, sociolinguistic influences, and environmental factors. The systematic changes and recognizable patterns highlight the dialect's adaptability and evolution. This study not only underscores the phonological richness of Multani Punjabi but also contributes to a deeper understanding of dialectal variation and language change within the Punjabi linguistic landscape.

References

Alizai, K. (2021). A case study of Punjabi language diglossia and language shift in Baluchistan Province. *Pedagogika Społeczna Nova*, *1*(2), 163–181. https://doi.org/10.14746/psn.2021.2.10

Arslan, M. F., Mahmood, M. A., & Haroon, H. (2021). Highlighting the Sound Shift in Punjabi Language: A Corpus-based Descriptive Study. *Linguistic Forum*, *3*(1), 1–5. https://doi.org/10.53057/linfo/2021.3.1.1

Crystal, D. (2008). A Dictionary of Linguistics and Phonetics (6th ed.). Oxford: Blackwell.

David, Maya & Manan, Syed. (2020). Linguistic Diversity as a Problem vs Linguistic Diversity as a Resource: A critique of language policies in Pakistan. 2019 . In Research Mosaics of Language Studies in Asia. Differences and Diversity. University Science, Malaysia.

Eberhard, David & Simons, Gary & Fennig, Chuck. (2020). Ethnologue: Languages of the World, 23rd Edition.

Gershon, I. & Manning, Paul. (2014). Language and Gupta, V., Lehal, G.S. (2011). Preprocessing Phase of Punjabi Language Text Summarization. In: Singh, C., Singh Lehal, G., Sengupta, J., Sharma, D.V., Goyal, V. (eds) Information Systems for Indian Languages. ICISIL

2011. Communications in Computer and Information Science, vol 139. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-19403-0_43
Hasan, Wasim & Arshad Khan, Ali. (2021). Study of Lexical Variation between Dhani and Majhi Punjabi. Global Language Review. VI. 1-16. 10.31703/glr.2021(VI-IV).01.
http://dx.doi.org/10.1002/9781444302776
Labov, W. (1972) Sociolinguistic Patterns (Philadelphia, PA: University of Pennsylvania Press) pp. 43-54
Mahmood, R. & Hussain, Q. & Mahmood, Muhammad. (2011). Phonological adaptations of English words borrowed into Punjabi. 22. 234-245.
Malik, Sitender & Kumar, Satish & Reena,. (2012). Village Information System
Malik, Sitender & Kumar, Satish & Reena,. (2012). Village Information System – media. 10.1017/CBO9781139342872.026. of Research in Social Sciences. 2. 184.
Rafi, M. S. (2010). *Semantic variations of Punjabi Toneme*. https://ssrn.com/abstract=3827789
Thomason, Sarah. (2006). Language Change and Language Contact.
Trudgill, P. (1995). Sociolinguistics: An Introduction to Language and Society. Penguin Books.

Jadavpur Journal of Languages and Linguistics
ISSN: 2581-494X

Phonological and Morpho-Syntactic Development in Early Childhood: A Cross-Sectional Analysis of Jordanian Arabic

Hosam Shatnawi & Mamta[1]
University of Delhi, Delhi

ARTICLE INFO

Article history:
Received 09/10/2025
Accepted 17/10/2025

Keywords:
morpho-syntax,
code-mixing,
gender agreement,
Jordanian Arabic (JA),
phonological awareness

ABSTRACT

Jean Piaget (1960) thought that children's cognitive faculties must develop before language does. His theory states that children go through stages of language development that result from their expanding knowledge of the world. This study intends to analyze the language development of young children, particularly in terms of Jordanian Arabic (JA). This study examines linguistic development in acquiring morphosyntax, Code-Mixing, phonological awareness, and gender agreement. There are a total 20 participants in the age groups of 3.5 to 7 years. The tests are conducted to analyze the language development in children. The research study looks at the responses of the children after watching a video of cats and dogs eating, drinking, and barking/meowing by the children. Data analysis uses both qualitative and quantitative techniques. The purpose of the study is to analyze the speech of children after watching video clips of cats and dogs eating, drinking and barking/meowing.

Literature Review

Arabic is a Semitic language spoken by millions. It is an important language in the Middle East and North Africa (Holes, 1995; Procházka, 2006; Owens, 2013; de Swaan 1998, 2001). In Arabic, the difference between Modern Standard Arabic (MSA), which is the written form of Arabic and colloquial dialects, is crucial. These dialects can differ substantially in how they sound, how they are spoken, and how they are constructed (Holes, 1995).

This study will discuss the specific features and development patterns of Jordanian Spoken Arabic. Studies suggest that children show chronological progression of language. At age 3 or 4, kids tend to simplify sound production by processes of syncope (deletion of a vowel), apocope (deletion of final sound) and consonant substitution (Daana, 2009; Sakarna, 1999). As children get older they show greater phonetic accuracy and mastery of complex grammatical structures.

In the area of gender agreement, there are complex rules governing how nouns, verbs and adjectives agree in gender and number. According to research done on Arabic dialects, younger children may have trouble with such complex agreements (Moawad, 2006, Ravid & Farah, 1999). Additionally, code-mixing can be insightful in understanding JSA's development, likely influenced by vocabulary acquisition and exposure to the other languages being used during the interactions.

A crucial aspect to note is that JA includes sub-dialects such as Bedouin and urban Ammani dialects (Cleveland, 1963). These dialects have minor differences in their sounds. As

[1] Corresponding Author Email ID: jnu.mamta@gmail.com

an illustration, the voiced interdental fricative /ð/ is present in the Bedouin dialects but replaced by the /d/ phoneme in the urban dialects (Sakarna, 1999). For this reason, we must examine the JSA acquisition closely.

Investigation into Arabic diglossia where MSA and spoken dialects coexist shows that children can perform better in MSA tasks when they are related to their spoken dialect, suggesting that MSA and dialect are not acquired the same way (Khamis-Dakwar, 2012).

Research Aims

1. Analyze Phonological Awareness and Accuracy: To determine how phonological awareness and accuracy in Jordanian Spoken Arabic (JA) change with age and to identify any age-specific patterns in sound production.
2. Understand Gender Agreement Rules: To investigate when children learn the complex rules of gender agreement in JA and to explore any differences in the development and problems faced by girls and boys.
3. Assess the Impact of Code-Mixing: To evaluate how significantly code-mixing with English influences the development of JA.

Methodology

Research design

A descriptive cross-sectional research design is adopted. This design enables researchers to study children's acquisition of specific language skills across various age levels. Descriptive research is the observation and description of communication that does not manipulate anything. It shows us how language abilities are naturally developed.

Sampling and Population

This study focuses on young children who speak Arabic in Jordan. The sampling will be purposive and 20 participants will be selected ensuring equal distribution of boys and girls from the age group of 3.5 to 7 years. To better understand the children's speech, these will be transcribed using a computer software named Praat. It will help them review the phonological production in detail. Through an analysis of their capacity to manipulate and reflect on the sounds of their language, I will gain insight into their developing linguistic competence in Jordanian Arabic.

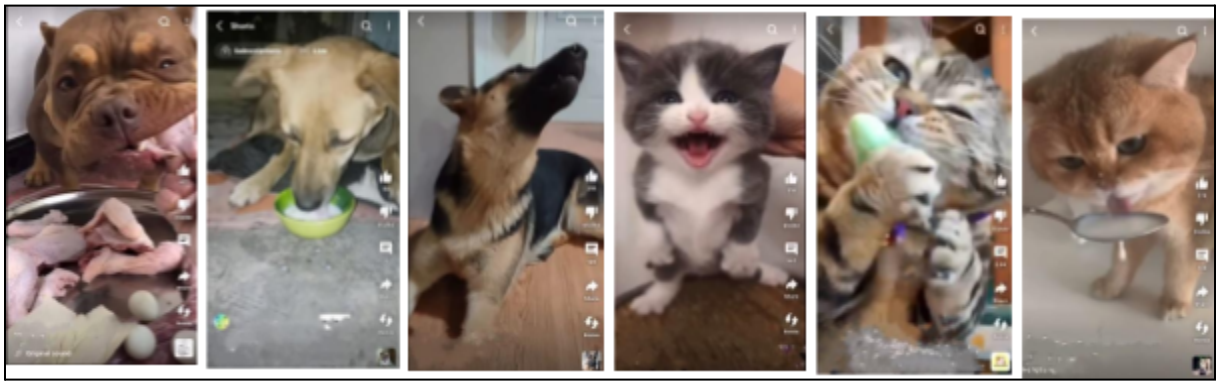
Data Collection

The present study used the natural speech of Jordanian Arabic-speaking children to collect data. Young children watched video stimuli showing cats and dogs doing various actions (eating, drinking, barking/meowing). Their responses when describing these actions were recorded, providing rich linguistic data for analysis. Additionally, I used Praat computer software to generate detailed phonetic transcriptions of their speech. This software enabled accurate analysis of sound productions in the child's speech.

Stimuli

This research utilized videos of cats and dogs doing different actions. These videos were made to be visually attractive so that children are encouraged to speak. After taking screenshots of the videos, capturing the moments of actions for the analysis. These screenshots show cats and dogs

doing common day-to-day activities like eating and drinking (Figures 1, 2, 5, 6). The stimuli also include dog barking (Figure 3) and cat mewing (Figure 4), which are specific to the species to



give the children plenty of ideas, the selected actions were quite varied. This way, we could see how children would describe actions that are both common and unusual in Jordanian Arabic.

Data Analysis

This research used a mixed approach in analyzing the data with a qualitative and quantitative perspective. It will provide a comprehensive analysis of the linguistic development of Jordanian Arabic-speaking children. We created detailed phonetic transcriptions of the speech samples of these children using the Praat computer software. By combining the quantitative and qualitative findings it would add another aspect to the research about the literacy level of Jordanian Arabic-speaking children.

Ethical Considerations

We strictly adhere to ethical considerations to ensure the participant's goodness, welfare, and rights. Before the research started, we got consent from all the children's guardians in charge of the child.

Review of Arabic

Arabic is a language that people in the Middle East and North Africa speak. It is the language of Islam too. This has made Arabic an official United Nations language (Holes, 1995; Procházka, 2006; Owens, 2013; de Swaan 1998, 2001). It should be noted that there is a difference between MSA (Modern Standard Arabic), used mostly in formal situations, and the various spoken dialects. These dialects may not be mutually intelligible with each other although they share a common base (Holes, 1995).

Jordanian Spoken Arabic (JSA)

This research investigates Jordanian-spoken Arabic (JSA) and the characteristics relevant to my study of JSA. Within Jordan, there are urban speakers, rural speakers, and Bedouin speakers (Cleveland, 1963; Suleiman, 1993). My research looks mainly at the urban Ammani and the Bedouin dialect.

Phonemic Inventory of Jordanian Spoken Arabic (JSA)

Jordan has diverse Arabic dialects. According to Cleveland (1963), four basic dialect labels have been identified based on socioeconomic/geographic (living place). Bedouin, rural, Palestinian (Jerusalem), and urban. The study will deal specifically with the urban Ammani and Bedouin

dialects of Jordan.

		Labial	Denti-alveolar		Palatal	Velar	Pharyngeal	Glottal
			plain	emphatic				
Nasal		m – م	n – ن					
Occlusive	voiceless		t – ت	tˤ – ط		k – ك		ʔ – ء ق
	voiced	b – ب	d – د	dˤ – ض				
Fricative	voiceless	f – ف	s – س ث	sˤ – ص	ʃ – ش	x – خ	ħ – ح	h – n
	voiced		z – ز ذ	zˤ – ظ	ʒ – ج	ɣ – غ	ʕ – ع	
Trill / Tap			r – ر					
Approximant			l – ل	(ɬ)	j – ي	w – و		

Table 1: *Consonantal Phonemic System of Urban Ammani Dialect*

Manner/ Place	Bilabial	Labio-dental	Inter-dental	Dental	Alveolar	Palato-alveolar	Palatal	Velar	Pharyngeal	Uvular	Glottal
stop	b			t/tˤ d				K g		q	ʔ
nasal	m				n						
flap					r						
affricate						tʃ dʒ					
fricative		f	θ ð/ðˤ		s/sˤ z	ʃ			ħ/ʕ	x ɣ	h
lateral					l						
glide	w						j				

Table 2: *Consonantal Phonemic System of Bedouin Dialect*

The tables show that the Bedouin dialect consists of 29 consonants while the urban dialect consists of 26. While some sounds are common to both dialects (i.e. /d/, /b/, /m/), others are interestingly different. The voiced interdental fricative /ð/ is one example. This sound is pronounced /ð/ in the dialect of the Bedouins, but /d/ in the urban dialect (Bedouin /ðura/, urban /dura/, both meaning 'corn').

In addition, the Bedouin dialect has a phoneme /ðˤ/. But in urban dialect, we can alternatively realize it as either /dˤ/ or /zˤ/, e.g., Bedouin /ðˤill/ vs. urban /dˤill/ or /zˤill/ (all meaning ‘shadow). Another difference is the voiceless interdental fricative /θ/, found in Bedouin dialects but realized as the plosive /t/ in urban dialects (e.g. Bedouin /kθi:r/ vs. urban /kti:r/, both meaning ‘a lot’).

In Bedouin, /t/ stays as a distinct phoneme. In addition, the Bedouin dialects have the alveopalatal affricate /tʃ/ as an allophone of /k/. For example, Bedouin /aħibbitʃ/ I love you [to female] has /k/ realized as /tʃ/, while the urban dialect has baħibbik. The fifth difference incorporates the consonant /dʒ/. When referring to Bedouin dialects, this consonant is

pronounced as /dʒ/, while urban dialects will realize this sound as /ʒ/. For example, the Bedouin form /dʒa:dʒ/ contrasts with the urban form /ʒa:ʒ/, which both have the same meaning ‘chicken’. Ultimately, the sound /g/ only appears in Bedouin dialects and is /ʔ/ in urban dialects. This is clear in the word /ga:l/ (meaning ‘he said’) in the Bedouin dialects, realized as /ʔa:l/ in the urban dialect (Al-Malahmeh, D. A. (2015)).

Morpho-phonology

The term “al-” means “the” in Arabic and gives nouns and adjectives a definitive purpose. Interestingly, when “al-” precedes certain nouns or adjectives, it assimilates with the first letter. This means the first sound changes to incorporate the /l/ sound of "al-".

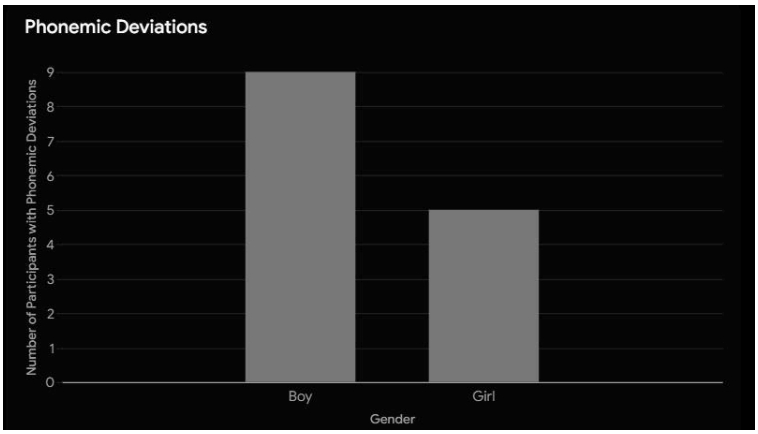
Consonants That Cause This Assimilation Are Known As Sun Consonants They include: /t/, /θ/ (th), /ð/ (th), /ðˤ/, /ʃ/, /r/, /z/, /s/, /d/son/, /l/, /n/, /tˤ/, /sˤ/, /dˤ/. For example, “al-ʃams” (“the sun”) is pronounced /aʃ-ʃams/. Here, the /l/ sound is assimilated into the /ʃ/. The other consonants are "moon letters" that do not cause the assimilation of the definite article. For instance, "Al-Quran" ("The holy Quran") will not be changed.

Morpho-syntax

In JSA, a suffix (-e:n) is added to the original words to form duals irrespective of case and gender. For instance, /binte:n/ “two girls” and /walade:n/ “two boys”. Pluralization occurs by adding suffixes (this is a “sound plural”; the marker for the feminine is –a:t, for the masculine is –i:n) or interdigitation ("broken plural"), the latter involving complex changes to a noun's root (Holes, 1995). Broken plurals, which have distinct meanings, are called these terms. For example, /beit/ ‘house/verse’ has plurals /bju:t/ ‘houses’ or /ʔabja:t/ ‘verses of poetry’ (Holes, 1995). Interestingly, numbers in JSA take the masculine form when they are used before a noun irrespective of whether the noun is masculine or feminine (e.g. /tlat bana:t/ “three girls” Moawad, 2006; Al-Aqarbeh, 2011). Adjectives, however, agree with their subjects in gender and number. The basic word order of adjectives in JA is post-nominal with respect to nouns, and numeral usually is pre-nominal (and taking a feminine form when this is the case). In the end, the JA has a use of masculine (/inta/) and feminine (/inti/) pronouns for the second person singular (Holes, 1995; Naess, 2008).

Results and Findings

Phonemic Deviations in Jordanian Arabic Phonology from Collected Data



The data shows differences in the ratio of phonemic changes according to gender. Moreover, boys appear to display a marginally increased frequency of phonemic deviations compared to girls.

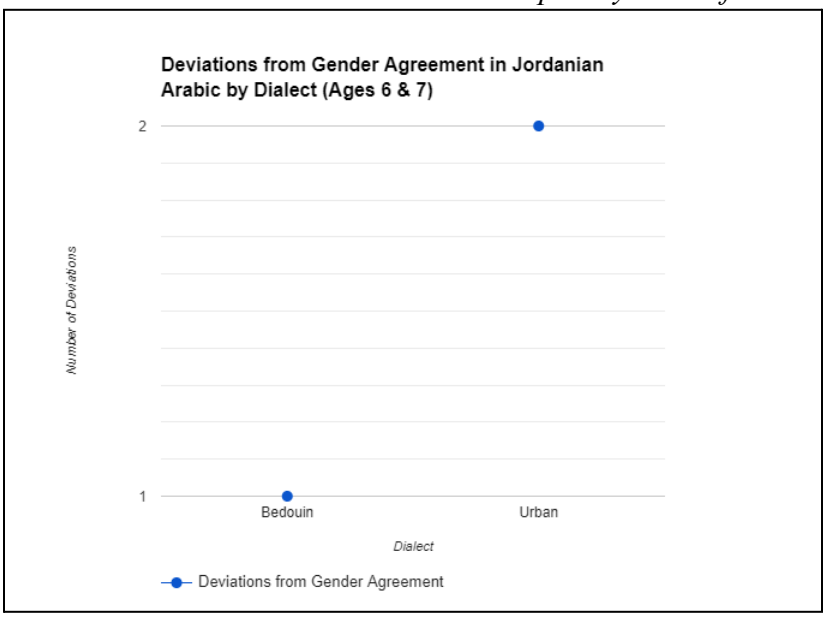
Analysing their Response

Case number	Age/Gender /Dialect	Transcribed Speech	Lexical Items	Phonemic Deviations in the lexical Items	Gloss
1.	7/ Boy/ Urban	*/btʃrab/*	*/btʃrab/*	Deletion of the consonant /r/	Drink
2.	3.5/ Boy/ Urban	*/algild/*	*/algird/*	Substitution in the mid-position consonant /r/	Monkey
3.	6/ Girl/ Urban	*/kab/*	*/kalb/*	Deletion of /l/ (Syncope)	Dog
4.	6/ Boyl/ Urban	*/tiθab/*	*/tiʃrab/*	Substitution of /ʃ/ and deletion of/r/	Drink
5.	7/Boy/Urban	*/btiʃʃb/*	*/btiʃrab/*	Deletion in /a/, Gemination of /ʃ/ and deletion in /r	Drink
6.	6/Boy/Urban	*/btihawi/*	*/bitʕawi/*	Substitution in the mid-position consonant /ʕ	Bark
7.	7/Girl/Urban	*/ga:ʕd/*	*/ga:ʕdah/*	Deletion in /ah/	Sitting
8.	6/Boy/Urban	*/ʕawi/*	*/bʕawi/*	/Procope in /b/	Bark
9.	6/Boy/Urban	*/ʔowdil/*	*/towkil/*	Consonant substitution of /t/ with the glottal stop	Eat

				/ʔ/ Substitution of the voiceless consonant /k/ to the voiced consonant /d/	
10.	6/Boy/Urban	*/btah/*	*/bisah/*	Syncope of the Med-vowel /i/ and the fricative consonant /s/ is replaced by the stop consonant /t/	Cat
11.	3.5/Boy/ Urban	*/laʔba/*	*/laħma/*	The voiceless pharyngeal fricative /ħ/ is replaced by a glottal stop /ʔ/ and Consonant substitution of /m/ with /b/	Meat

The data shows the main phonological changes in the child from different ages, genders, and dialects. Most often this includes deletion of consonants, substituting consonants, and deletion of vowels/consonants. This data is just a sample and does not represent the whole phonological shifts. The ratio of phonemic changes in whole data varies according to gender. Boys seem to show a slightly higher frequency of phonemic errors than girls.

Morpho-Syntactic deviations in Jordanian Arabic Morpho-Syntactic from Collected Data



The data shows differences in the ratio of Gender Agreement deviation according to Number of deviation and dialect. According to the findings, there was very little deviation, as the Bedouin dialect had one and the Urban dialect had two.

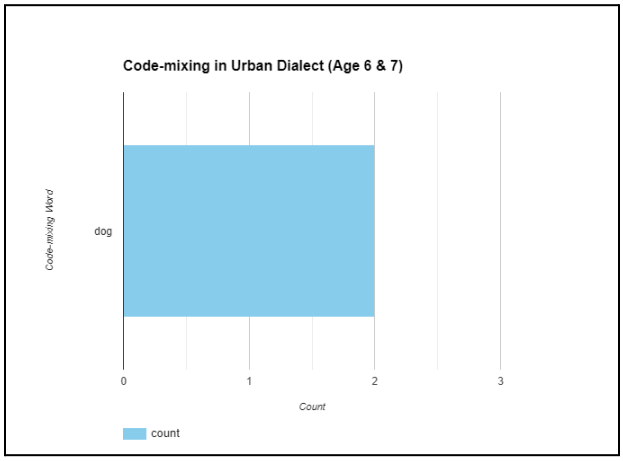
Analysing their response

Case Number	Age/Gender/Dialect	Transcribed Speech	Deviation Type	Explanation
1.	6 yrs. / Girl / Bedouin	*/bisah gaʕd ta:kul χja:r wbrukali:/*	No gender agreement	The verb "gaʕid" (masculine) should be "gaʕdah" (feminine) to agree with the feminine noun "bisah".
2.	7 yrs. / Girl / Urban	*/ʃuft bisah ga:ʕd btiʃrab ħali:b/*	No gender agreement	The verb "gaʕid" (masculine) should be "gaʕdah" (feminine) to agree with the feminine noun "bisah".
3.	7 yrs. / Girl / Urban	*/ʃuft bisah qa:ʕid btʕawi/*	No gender agreement	Similar to the previous example, the verb "qa:ʕid" (masculine) should change to "qa:ʕdah" (feminine) for agreement.

The table above shows how young Arabic-speaking girls in Jordan disagree in gender and consistently apply an incorrect masculine verb form instead of the correct feminine form when a feminine noun is involved. Interestingly, boys showed no similarity deviations, suggesting that the acquisition of these rules could involve a gender-specific pattern.

Code-Mixing Findings and Variations in Jordanian Arabic from Collected Data

Instances of code-mixing with English were observed in the Urban dialect among 6 and 7-year-old boys.



Analysing their Response

Age/Gender	Dialect	Transcribed Speech	Gloss	Code-Mixing Elements
7 yrs./Boy	Urban	*/ʃuft du:g wa:gif/*	I saw a dog standing	Code-mixing with "dog"
6 yrs./Boy	Urban	*/ʃuft du:g gaʕd bukil/*	I saw a dog standing	Code-mixing with "dog"

Two cases of code-mixing are recorded among children belonging to the Urban dialect. The English word “dog” is inserted in both sentences in Arabic. This suggests that the child may mix codes due to not knowing the word and the word being part of their everyday speech.

Conclusion

The research has revealed important information about the development of Jordanian Arabic (JA) in children aged 3.5-7 years. Our finding shows age, gender and dialect-related sound and structure differences across speakers. Younger children had more issues with phonological deviance that tended to reduce as the children aged, demonstrating progress in phonetic accuracy and phonological awareness. It was also noticed that girls made more mistakes in gender agreement compared to boys.

Code-mixing with English was influential, especially urban boys. It has resulted in the influence of two languages on JA. Also, a comparison of Bedouin and Urban 2 dialects showed they have phonemic and syntactic differences too. The findings underscore how complex learning languages is and how sociolinguistic factors affect the linguistic learning of children. They help us understand and learn how children in a diglossic society acquire the linguistic structures of their mother tongue.

References

Procházka, S. (2006). Arabic. In K. Brown (Ed.), Encyclopedia of language and linguistics (pp. 423-431). Elsevier.

De Swaan, A. (1998). A political sociology of the world language system (1): The dynamics of language spread. Language Problems & Language Planning, 22(1), 63-75.

De Swaan, A. D. (2001). Words of the world. Cambridge: Polity Press.

Holes, C. (1995). Modern Arabic: Structures, functions, and varieties. Georgetown University Press.

Daana, H. A. (2009). The development of consonant clusters, stress and plural nouns in Jordanian Arabic child language. (Unpublished doctoral dissertation). The University of Essex.

Sakarna, A. K. (1999). Phonological aspects of Abady Arabic, a bedouin Jordanian dialect. (Unpublished doctoral dissertation). University of Wisconsin Madison.

Moawad, R. A. (2006). The acquisition of the Arabic gender and number systems. (Unpublished doctoral dissertation). University of Wales.

Ravid, D., & Farah, R. (1999). Learning about noun plurals in early Palestinian Arabic. First Language, 19(56), 187-206.

Cleveland, R. L. (1963). A classification for the Arabic dialects of Jordan. Bulletin of the American Schools of Oriental Research, 56-63.

Khamis-Dakwar, R. (2011). Early Acquisition of SVO and VSO word orders in Palestinian Colloquial Arabic. In Perspectives on Arabic Linguistics: Papers from the annual symposia on Arabic Linguistics. Volume XXII-XXIII: College Park, Maryland, 2008 and Milwaukee, Wisconsin, 2009 (Vol. 317, p. 281). John Benjamins Publishing.

Khamis-Dakwar, R., Froud, K., & Gordon, P. (2012). Acquiring diglossia: Mutual influences of formal and colloquial Arabic on children's grammaticality judgments. Journal of Child Language, 39(01), 61-89.

Boersma, Paul & Weenink, David (2024). Praat: doing phonetics by computer [Computer program]. Version 6.4.06, retrieved 25 February 2024 from http://www.praat.org/

Dima Awadh Al-Malahmeh (2015) An investigation of the influence of sociolinguistic factors on children's first language in Jordan.

Al-Aqarbeh, R. N. (2011). Finiteness in Jordanian Arabic: a semantic and morphosyntactic approach. (Unpublished doctoral dissertation). Kansas University, Lawrence, KS.

Næss, U. G. (2008). "Gulf Pidgin Arabic": Individual strategies or a structured variety?. (Unpublished doctoral dissertation). University of Oslo.

. Piaget, J. and B. Inhelder, 1960. The growth of logical thinking from childhood to adolescence, Routledge and Kegan Paul, London.

Jadavpur Journal of Languages and Linguistics
ISSN: 2581-494X

A Comparative Study of Compound Word Formation in Assamese & Bodo Languages

Banani Das[1]
The Assam Royal Global University, Assam

ARTICLE INFO

Article history:
Received 15/09/2025
Accepted 06/10/2025

Keywords:
comparative language,
Bodo,
Assamese,
morphology,
word formation,
compounding

ABSTRACT

In the realm of linguistics, word formation processes play a pivotal role in understanding how languages evolve and adapt. One such process, compounding, serves as a fundamental mechanism for creating new words by combining two or more existing words. This research paper delves into the intricate workings of compounding within the linguistic frameworks of Assamese and Bodo, two prominent languages spoken in the northeastern region of India. Through a comparative analysis of compounding in Assamese and Bodo, this research paper seeks to shed light on the similarities and differences in how these languages employ compounding as a word formation process. By examining a diverse range of lexical creations and semantic extensions within compound words, this study aims to provide valuable insights into the linguistic mechanisms at play in these languages, offering a deeper understanding of how compounding shapes the lexicons of Assamese and Bodo speakers.

Introduction

Morphology as a sub-discipline of linguistics deals with the study of the internal structures of words and how they are formed by the smaller pieces (Booij,2005). The exploration of morphology extensively delves into word formation processes, and the examination of compound word formation is a well-established subject. Since 1916, Saussure's theory of synchronic and diachronic linguistics has had a great impact on the field of modern linguistics, yet its closely related subject of lexicology was simply put aside and paid little attention to. Linguists either focused on synchronic linguistics as L. Bloomfield did or just adopted diachronic linguistics as Kozoid did. In 1957, Noam Chomsky had his "Syntactic Structure" published, and syntax became a focus at that time while other structuralists went on with their studies on phonetics and morphology. In 1960, H. Marchand published his influential book "The Categories and Types of Present-Day English Word-formation", in which he disregarded synchronic and diachronic approach to word study. During the 1980s, linguists conducted their studies on word formation from different points of views: phonetic, synchronic and semantic respectively. More linguists intended to find a universal grammar through the studies on the formation of words (Laurie Bauer, 1983:2-6).

The origins of studying compound words trace back to 400 B.C., when Indian linguist Panini meticulously detailed the formation of Sanskrit words in his renowned grammar treatise, the "Eight Chapter Book." Panini's grammar, as highlighted by Kiparshy (2008), has received universal acclaim for its comprehensive analysis of Sanskrit. Empiricist Bloomfield also lauded

[1] Corresponding Author Email ID: bananidas.hazarika@rgi.edu.in

Panini's grammar for its simplicity and synchronic viewpoint. Panini extensively covered word formation processes in Sanskrit, including the intricacies of compound word formation.
In the 2nd Century B.C., Patanjali further observed and documented the formation of compounds in language. Their studies led to the development of mechanisms that are still utilized as foundational tools in word analysis by numerous linguists. Noteworthy compound types such as Bahuvrihi (exocentric compound), tatpurusa (determinative), and dvandva (coordinate compound) emerged from their work. These mechanisms continue to be actively applied in the study of compound formation within languages, with contemporary linguists frequently referencing and discussing them. The enduring relevance and applicability of Panini and Patanjali's insights underscore the enduring impact of their work on the field of linguistics.

Compounding, a prevalent word formation process in linguistics, involves combining two or more words to create a new lexical unit. In languages like Assamese and Bodo, compounding plays a significant role in expanding vocabulary and expressing complex ideas. The structure of compound words in Assamese follows specific patterns, often reflecting cultural concepts and societal norms. Similarly, in Bodo, compounding serves as a means of conveying nuanced meanings and reflecting the unique worldview of its speakers. By examining the structural characteristics and semantic implications of compound words in Assamese and Bodo, researchers gain valuable insights into the underlying mechanisms of word formation in these languages. The study of compounding within these linguistic contexts provides a deeper understanding of how speakers construct meaning through the strategic combination of existing lexical elements, enriching the lexicons of Assamese and Bodo speakers.

Word Formation

The process of word formation in languages stands as a critical aspect of linguistic analysis, shedding light on how words are created and structured to convey meaning effectively. Word formation processes vary across languages, encompassing a wide array of strategies such as compounding, derivation, blending, and conversion. Through an in-depth examination of these processes, this study aims to unravel the underlying principles that govern how new words emerge within linguistic systems. By investigating the structural patterns, semantic implications, and cultural influences embedded in word formation, this research seeks to provide valuable insights into the dynamic nature of language evolution.
The study of word formation extends beyond mere linguistic analysis, offering a window into the socio-cultural landscapes and historical developments of different language communities. By tracing the etymology of words and dissecting the morphological intricacies of word creation, researchers can uncover a wealth of information about the evolution and diversification of languages over time.

Assamese, a language with deep historical roots and a rich literary tradition, showcases a complex system of compounding where elements merge to form cohesive lexical units. By exploring the nuances of compounding in Assamese, this study aims to unravel the structural patterns, semantic implications, and cultural influences that shape the formation of compound words in this language. On the other hand, Bodo, an indigenous language spoken by the Bodo people, offers a unique perspective on compounding, showcasing distinct features and variations compared to Assamese. By analyzing compounding in Bodo, this research endeavors to highlight the structural divergences, semantic shifts, and sociocultural aspects that underpin the creation and usage of compound words in this linguistic context.

Compounding

Compounding is the process of bringing together two or more words into a single entity that has one meaning of the word. According to Matthews (1974: 82) compounding is a process by which a compound lexeme is derived from two or more simpler lexemes. It is a combination from form lexemes to not simply form.

In morphology compound is realized as a word formed by the combination of two or more independent words. In the opinion of David Crystal (2008) compound refers to a linguistic unit which is composed of elements which function independently in other circumstances.' J.Toman (1992) is of the opinion that 'compound is a morphologically complex word containing at least two elements, which can otherwise occur as a free form'. Straightly speaking compound is a derived word formed by the combination of morphological units (elements) and it functions as an independent word in the language.

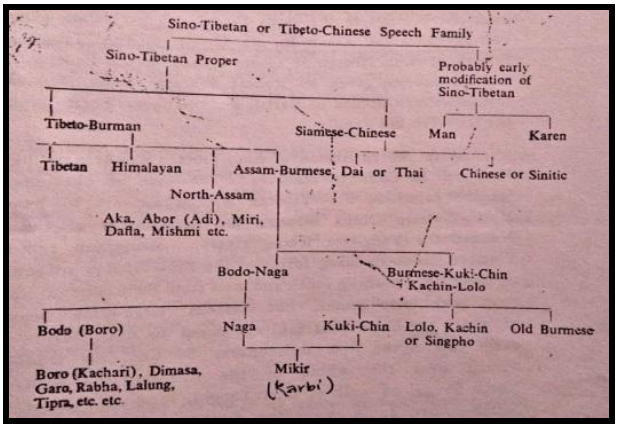
The new words which are made by the compounding process are battleground, audio dub, aussieland, batchmate, hackboat, etc. For detailed data, see appendix 1. The word 'battleground' is made from (battle+ground), audio dub (audio+dub), aussieland (Aussie+land), batchmate (batch+mate),hackboat (hack+boat). From the explanation above, it can be seen that those new words are coming by combining two words.

Background Bodo and Assamese languages

Bodo, a language of Bodo-Garo group of Sino-Tibetan Speech family of North-East India.
Bodoland, officially the Bodoland Territorial Region, is an autonomous region in Assam, Northeast India. It is made up of four districts on the north bank of the Brahmaputra river below the foothills of Bhutan and Arunachal Pradesh (Bhattacharya, 2007).

Assamese is a language of the Indo-Aryan language group. Although both the languages belong to different language families they show certain similarities in their language structure. In the language among many processes of lexical composition compound formation bears a significant role and exhibits its productivity prominently. But this interesting compound formation of the languages is so far neglected in the linguistic study of the language. Therefore, this paper intends to provide a new insight to the compound formation of the languages and focussing on the similarities in terms of compounding process of word formation.

The speakers of these two languages have been living together in the same geographical area from earlier periods as good neighbors. So the Bodo speakers have incorporated some elements from the Indo-Aryan Assamese speakers and the Assamese speakers have incorporated some elements from the Tibeto-Burman Group of Bodo speakers. These mutual influences are basically seen in the case of Phonology, Morphology, Syntax , Lexis of vocabulary and other sociocultural habits, and it helps to develop the socio-cultural integration in the region.



Furthermore, Mustafa et al.(2015) stated that word formation processes which are very common to use by Malaysian Facebook users are abbreviation, blending, and the use of emoticons while they communicate on Facebook every day. Then, Wei and Wenyu (2014) said, "The most frequently occurring word-formation process of netspeak neologisms is compounding, subsequently, blending, affixation, old words with new meaning, acronyms, conversion, and clipping".

Objectives

Word-formation is the study of words, dealing with the construction or formation rules of words in a certain language. Though the Assamese and Bodo language are sprung out from the different roots or sources, yet they possess some common features. The aim and objective of this proposed paper is to put up a comparative picture of the similarities of the word formation processes of Assamese and Bodo languages of the state of Assam.

Hence, the paper is an attempt to discuss the word formation process of both the languages in a comparative manner. The aims and objectives of this paper are-

1. To study the Compound word formation process in Bodo and Assamese.
2. Analyses various ways of compound word formation in Assamese and Bodo languages so that similarities and differences are found between the two languages.

Through comparison, Axomiya and Bodo learners could have a better way of acquiring either or both of the two languages concerned.

Methodology

This study was designed as qualitative research because it presents the data by describing the processes of word formation in creating new words. As per methodology is concerned this paper is primarily based on secondary information. The secondary information is collected from different sources like books, journals, souvenirs, research articles etc.

In achieving the objectives, the researcher uses descriptive qualitative as the type of the research. The data are in the form of compound words. In collecting the data the writer uses documentation technique selecting and collecting the data systematically in accordance with the types and the meaning of compound words.

Discussion

Compounding is a productive morphological process of word formation of Assamese and Bodo language. The common compound words of both the languages can be discussed classifying into the following compounding types.

1. Pure compound
2. Hybrid compound
3. Borrowed compound
4. Copulative compound
5. Bahuvrihi (exocentric) compound

The sections below discuss the different types of compound formation in both the languages. It is seen that in both the languages compounds are formed by adding two different roots. Both the two roots carry some meaning. Sometimes the meaning is inherited by one of the two roots and sometimes they denote a different meaning.

Endocentric Compounds in Bodo and Assamese:

A compound will be called "endocentric" if it contains a "head" morpheme that determines its morphosyntactic features and general semantic type. The head of the compound is that root which determines the category of, or is the main referent of the compound. For instance, in Liangmai language (Naga) kəbui-dui 'milk' is an endocentric compound consisting of two elements, namely kəbui 'cow' and dui 'water' (Damai, 2019). In Endocentric compounds the meanings are inherited by one of the two words which are combined together to create a new lexeme. In this compound dui meaning 'water/milk' functions as the head and kəbui 'cow' functions as the attribute of dui (water). In Endocentric Compounds, the referent of the compound is always the same as the referent of its head. For example in the word "flower book" 'book' is the head morpheme. Hence the compound is a noun (like book), and names a type of book.

4.1.a. Endocentric compounds in Boro

Head 1	Head 2	Head 3	Compound word	meaning
baɾ 'wind'	duɯi 'water'	sikʰla 'girl'	baɾduɯisikʰla	'weather goddess'
sikʰɾi 'butterfly'	sikʰla 'young girl'		sikʰɾi-sikʰla	'a young girl like a butterfly'

Table 1*: Endocentric compounds in Boro*

4.1.b. Endocentric compounds in Assamese

Head 1	Head 2	Head 3	Compound word	meaning
বিয়া (marriage)	ঘৰ (house)		বিয়াঘৰ = /biagʰɒɾ	A marriage hall/house
ৰভা (pandal)	তল (under)		ৰভাতল = /ɾɒvatɒl	Function pandal
kʰeti (cultivation)	pɔtʰar (field)		kʰetipɔtʰar	Paddyfield

Table 2*: Endocentric compounds in Assamese*

Headedness in Endocentric compounds:

In syntax, the notion head characterizes the dominant member in an asymmetrical relationship within a construction (cf. Croft 2001: 241). The notion was explicitly applied to all

morphological constructs in the early 1980s by Williams (1981) and Selkirk (1982) and defended strongly (cf. Štekauer 2000; Di Sciullo and Williams 1987). However, not all linguists share the view that the notion extends naturally to all morphological constructions, especially affix-derived words (cf. Bauer 1990; Zwicky 1985). For compounding, though, the 8 notion head has existed longer and has had a central role in its study (cf. Scalise and Fábregas 2010). (Appah, 2017, p.7)

The classification of compounds may be based on the presence of a head constituent. It may also be based on the word class of the head (yielding nominal compound, verb(al) compound, adjectival compound, etc.) or the position of the head in the compound, (yielding left-headed/right-headed compounds). However, there appears to be a near-universal preference for right-headedness (Appah, 2017)

This motivated Williams (1981: 248) to propose what was initially presented as an absolute universal, the so-called right-hand head rule which he formulated as: "[i]n morphology, we define the head of a morphologically complex word to be the righthand member of that word". Later research showed that left-headedness also abounds, especially in the Romance languages which have predominantly left-headed compounds (cf. Scalise 1984). Therefore, it was suggested that the head position could be a parametric feature, so that the morphology of a language is either left-headed or right headed, depending on the language (Lieber 1992; Scalise 1992; Selkirk 1982). However, this idea of parametric variation in headedness has been shown not to be sustainable because in many languages, including Mandarin Chinese and Vietnamese, there are both left-headed and right-headed compounds (Ceccagno and Scalise 2006; Ceccagno and Basciano 2009; Packard 2000). A language like Nizaa (BenueCongo, Cameroon) even has an almost equal number of right-headed and left-headed N-N compounds (Pepper 2010). Appah (2013a) has shown that Akan endocentric N-N compounds are mostly right-headed, but there are left-headed and dual-headed N-N compounds as well. The semantic head is the constituent that shares its lexical conceptual information with the whole compound, so that the compound will be a hyponym of the semantic head (Scalise and Guevara 2006: 190). That is, the class of elements denoted by the compound will be a subset of the class of elements denoted by the head of the compound. For example, a table mat is a type of mat, the head. The formal head is the constituent that shares its formal features (lexical category and subcategorization frame) with the whole compound, so that the whole compound has the same distributional properties as its formal head (Scalise and Guevara 2006: 190). That is, mat can occur where table mat is expected to occur without a drastic change in the meaning of the construction, but the same cannot be said about the non-head constituent table occurring alone where table mat is expected. Booij (2012: 79) sums this up nicely in the observation that "headedness of a compound is not only relevant for its formal properties, but also for its semantic interpretation'.

4.2.a. Right-Headed compounds in Boro

Head 1	Head 2	Compound word	Meaning
bɔɾɔ 'Bodo'	ma 'mother'	bɔɾɔma	mother of the Bodos'
lai 'leaf'	pʰaŋ 'tree'	laipʰaŋ	'leafy plant'
sɯima 'dog'	gɔn 'stick'	sɯima gɔn	a post supporting wall'
dɯikʰɔŋ 'riverine'	kʰuli 'place'	dɯikʰɔŋ-kʰuli	'riverbank'

***Table 3** : Endocentric Right-headed compounds in Boro language*

4.2.b. Right-headed Compounds in Assamese

Assamese language also shows the presence of right-headed compounds as in the example /raz/ means king or kingly and /xɔbha/ means meeting. But the word does not specify anything about king or kingly it indicates a special king of meeting. And the internal meaning carried by the word which lies on the right side of the word. In the case of a right-headed <noun + noun> compound the first root modifies the second root as shown in examples below. The resulting meaning of the noun + noun compound is always a noun.

Head 1	**Head 2**	**Compound word**	**Meaning**
pitɔl ‘brass’	soku ‘eye’	Pitolsokua	One who’s eyes are brass colored
Goz ‘elephant’	mukʰ ‘face’	gozmukʰ	One who’s face is alike elephant
xadʱu ‘honest,saint’	purux ‘person’	xadʱupurux	An honest person
Raz ‘kingly/king	xɔbʱa ‘meeting’	razxɔbʱa	A royal court
zɔgɔt ‘universe’	nath ‘Lord, protector’	zɔgɔnnath	God
puspɔ ‘flower’	anjɔli ‘offering’	puspanjɔli	flower offering

Table 4: *Endocentric Right-headed compounds in Assamese language*

4.3.a. Endocentric Left- headed compounds in Boro

Head 1	**Head 2**	**Compound Word**	**Meaning**
ɔr (fire)	salai (tongue)	ɔrsalai	Flame
nɔ (house)	gɯthaŋ (green)	gɯthaŋnɔ	Greenhouse
khan (to comb)	zɔŋ (spear)	khanzɔŋ	Comb

Table 5: *Endocentric left-headed compounds in Bodo*

4.3.b. Endocentric left-headed compounds in Assamese

Head 1	Head 2	Compound word	Meaning
kotʰa ‘speech’	upɔkɔtʰɔn ‘a story’	kotʰupɔkɔtʰɔn	A conversation’
debə ‘God’	rixi ‘saint’	debərxi	A God who is also a saint.
raza ‘king’	rixi ‘saint	razɔrxi	A king who leads an ascetical life’
purux ‘person’	xadʱu ‘honest, saint’	xadʱupurux	An honest person
mukʰɔ ‘face’	Sɔndrɔ ‘moon’	mukʰɔSɔndrɔ	A special kind of face like the moon.

Table 6: *Endocentric left-headed compounds in Assamese*

Bahuvrihi compound (Exocentric)

In the realm of compound words, the concept of "exocentricity" arises when a compound lacks a pronounced head, with the understood head not overtly expressed. For instance, in the English compound pick-pocket, where the implied head is "person," denoting someone who picks pockets (Synder, 2001), the compound becomes exocentric. Bahuvrihi compounds, on the other hand, consist of two free morphemes with distinct meanings, yet the individual meanings of the

components are not preserved after compounding. Instead, these compounds signify a specific trait or characteristic possessed by an individual (Appah, 2017).

Exocentric compounds form a distinct entity, falling into a different category in terms of external distribution, making it challenging to deduce the compound's meaning solely from its constituent parts. Bauer (2008) emphasizes the presence of an unexpressed head carrying essential features of the compound's overall meaning, complicating the interpretation of exocentricity.

In Bauer's typology of exocentric compounds, bahuvrihi is a principal type. Indeed this Sanskrit expression is commonly used as a cover term for exocentric compounds (cf., inter alia, Bloomfield 1933; Katamba and Stonham 2006; Marchand 1969). Bahuvrihi is the type of 14 compounds in which the whole is not a hyponym of either constituent. Rather, it "expresses some salient facet of the denotatum" (Bauer 2008: 56).Bauer's analysis of exocentric compounds, notably bahuvrihi, positions it as a primary type. Bahuvrihi compounds, such as bahu vrihi meaning "much rich," exemplify this category by expressing a significant aspect of the denotatum without directly naming a person or entity (Bauer, 2008). This possessive interpretation of bahuvrihi compounds distinguishes them from other compound types, showcasing a nuanced approach to conveying specific attributes or qualities.

Regarding the English example red-eye which does not denote anything which possesses red eye but means "cheap whisky/overnight flight", Bauer (2010: 167) observes that the compound refers to that which causes red eye – the cheap whisky or overnight flight. Thus, from Bauer (2008, 2010), we may distinguish at least two types of bahuvrihi compounds – the possessive type and the causer type.

Furthermore, Bauer delineates two main types of bahuvrihi compounds—possessive and causer types. The possessive bahuvrihi compounds typically feature an adjective followed by a noun, highlighting the possessed element that characterizes a person or object (Bauer, 2010). This intricate analysis underscores the complexity and diversity of word formation processes, particularly in the context of exocentric compounds like bahuvrihi.

Some bahuvrihi compounds of Bodo and Assamese language have been illustrated in the following section.

4.4.a. Bahuvrihi (Exocentric) compounds in Boro

Head 1	Head 2	Compound word	Meaning
akhai (hand)	gɯlaɔ (long)	akhai-gɯlaɔ	Powerful
bikha (heart)	guar (wide)	bikha-guar	kind-hearted
daok^{h}a 'crow'	k^{h}amp^{h}ai 'flat seat'	daok^{h}a-k^{h}amp^{h}lai	a kind of creeper fruit'
sɯima 'dog'	sik^{h}aɾi 'hunter'	sɯma-sik^{h}aɾi	a kind of bee which makes its nest with clay'
lama 'road'	ɠidiŋ 'turn'	Lamdiŋ	name of a place in Assam

Table 7: *Exocentric compounds in Boro*

4.4.b. Exocentric compounds in Assamese (Jugalakshmi, 2020, p.39)

Head 1	**Head 2**	**Compound word**	**Meaning**
bina (a musical instrument)	pani (hand)	binapani	the goddess of learning and music
nila (blue)	ombor (cloth)	nilambor	the elder brother of Lord Krishna
podmo (lotus)	nab (naval)	podmonab	Lord Vishnu
nɔrɔ 'human'	xiŋhɔ 'lion'	nɔrɔxiŋhɔ	The lion headed the incarnation of Vishnu.
mrigɔ 'deer'	Trixa 'thirst' , 'ernest desire'	mrigɔtrixa	Mirage
rɔtna 'jewel'	Kor 'do'	rɔtnakor	the other name of sage Valmiki
mɔha 'uncle', 'very'	udɔi 'rise'	mɔhudoi	A respectful or formal term of address used for a man

Table 8: *Exocentric compounds in Assamese*

Copulative compounds (Hyphenated)

A copulative compound is that type of compound where two components are joined by copula but the copula (i.e.'or', 'and') does not occur in the compounding structure. Most of the copulative compound words are expressed by inserted hyphens in between their coordinate components. That is why such compounds may be termed as hyphenated compounds. In copulative compounds the constituents refer to properties of the same entity. For instance, a tuinman-chauffeur gardener-chauffeur is someone who is both a gardener and a chauffeur. Adjectival compounds of this type are, for example, doof-stom deaf-dumb and rood-wit-blauw red-white-blue. It is a type of compound in which both the words are equally powerful and have equal weightage.

The copulative compounds used in Bodo and Assamese are found to be of noun + noun, adjective + adjective, verb + verb structures. In Bodo, there are a good number of copulative compounds which denote collective meaning along with the sum of the elements participating in compounding. This type of compound can be said to be Collective Copulative Compound.

4.5.a. Copulative compounds in Boro

Head 1	**Head 2**	**Compound Word**	**Category**
Ha 'soil'	dɯi 'water'	ɟha-dɯi 'soil and water '	N +N
San 'day'	hɔr 'night'	san-hɔr 'day and night'	N +N
mɯzaŋ 'good'	gazri 'bad'	mɯzaŋ-gazri 'good and bad	Adj +Adj
undɯi 'small'	Geder 'big	undɯi-geder 'small and big'	Adj +Adj
ɾaŋ 'rupee'	ɾupʰa 'silver'	ɾaŋ- ɾupʰa 'property'	Adj +Adj
Za	Undu	za-undu 'eat and sleep'	V+V
thɯi 'do'	Thaŋ 'die'	thɯi-thaŋ 'do or die'	V+V

Table 9: *Copulative compounds in Boro*

4.5.b. Copulative compounds in Assamese

Head 1	Head 2	Compound Word	Category
Mama 'maternal uncle'	bʰagin 'nephew'	mama-bʰagin	N+N
Pita 'father'	putrɔ 'son'	pita-putrɔ	N+N
gɔhɔna (ornaments)	Pati (etc.)	gɔhɔna-pati	
Deh 'body'	Pran 'soul'	Deh – pran	N+N
kagɔz 'paper'	pɔtrɔ 'letter'	kagɔz- pɔtrɔ	N+N
Puja	Upaxona	Puja-upaxona 'cult'	N+N
ma:n	mɔrjyoda:	ma:n-mɔrjyoda: 'prestige'	N+N
ɾaŋ (colour)	ɾupʰ (beauty)	ɾaŋ-ɾupʰ 'beauty'	Adj +Adj
		alɔi-athoi 'Helplessness'	Adj +Adj
Era 'leave'	Dhora 'hold on'	Era-dhora	V + V
ma:r 'hit'	Dhor 'hold'	Ma:rdhor 'fight'	V+V
Ghuxona 'announcement'	pɔtrɔ 'letter	Ghuxona-pɔtrɔ 'placard'	V +N
Mila (to meet)	Pri:ti 'love'	Mila-pri:ti (to live in harmony)	V+Adj

Table 10: *Copulative compounds in Assamese*

Compound Noun

Compound nouns are extensively used in both languages. In the formation of compound nouns the words which are combined together belong to different word classes. The combination of words which is taking place in the formation of noun in the language can be determined by the following compounding processes-

4.6.a. Compound noun in Boro with N+N : N

The process of combining noun with noun also constitutes a good number of nouns of Bodo vocabulary.

Head 1	Head 2	Compound Word	Meaning
gaŋ (wing)	guma (grasshopper)	gaŋgɔma	one kind of winged grasshopper
sɔlɔ 'story'	khɔntha	sɔlɔ-khɔnthai	Epic
apha (foot)	khi (stool)	aphakhi	Footprint
hi (cloth)	sɔri (rag)	hisri	Rag
ɔr (fire)	salai (tongue)	ɔrsalai	Flame
megɔn (eye)	dɯi (water)	mɯdɯi	Tears
khuga (mouth)	dɯi (water)	khudɯi	Saliva

Table 11: *Compound Noun Compounds in Boro (N+N : N)*

4.6.b. Compound noun in Assamese with N+N : N

Head 1	Head 2	Compound Word	Meaning
nɔndɔ 'Lord Krishna'	utxɔb 'festival'	nɔndutxɔb	A festival for Lord Krishna
prɔsnɔ 'question'	uttɔr 'answer'	prɔsnuttɔr	Question-answer
zɔgɔt 'universe'	nath ''Lord, protector'	zɔgɔnnath	God
nɔrɔ 'human'	xiŋhɔ 'lion'	nɔrɔxiŋhɔ	The lion headed the incarnation of Vishnu.
bidya 'knowledge'	xagor 'ocean'	bidjaxagor	Knowledge like an ocean
bɔn 'forest'	bhuj 'feast'	bɔnbhuj	Picnic
tirthɔ 'holy place'	jatri 'one who travels'	tirthɔjatri	Pilgrim
tirthɔ 'holy place'	bhrɔmɔn 'journey'	tirthɔbhrɔmɔn	Pilgrimage
jɔl 'water'	dɔsyu 'bad person'	jɔl dɔsyu	Pirate
Jahaz 'ship'	Ghat 'bank'	Jahazghat	Port
xɔbha 'assembly'	Nripoti 'king'	xɔbhapoti	Spokesman, Chairman
Deb 'god'	Axur 'demon'		

Table 12: *Compound Noun Compounds in Assamese (N+N : N)*

Compound formation with Noun + Adjective: Noun word structure

Combination of noun and adjective is also one of the processes that is used in the formation of compound nouns in the language. But the nouns formed by this process are not abundant in number in the language.

4.7.a. Compound Noun in Boro with N+Adj : N

Head 1 Noun	Head 2 Adjective	Compound Noun	Meaning
nɔ 'house' – N	blaŋ 'wide open' – Adj	nɔblaŋ	Auditorium - N
nɔ (house)	gɯthaŋ (green)	gɯthaŋnɔ (greenhouse).	Greenhouse - N

Table 13: *Compound Noun Compounds in Boro (N+Adj : N)*

4.7.b. Compound Noun in Assamese with N+Adj : N

Head 1 (Noun)	Head 2 (Adjective)	Compound Noun	Meaning
rixi 'saint'	mɔha 'great'	mɔhɔrxi	A superior sage
zɔn 'person'	xɔt 'honest	xɔʤɔn	Honest person
purux 'person'	xadʱu 'honest, saint'	xadʱupurux	An honest person
Ha:t 'hand'	digʰɔl 'long'	Ha:tdigʰɔl	Powerful, influencial

Xima 'boundary'	Sari 'four'	Sarixima	Precinct 'boundary of a place'
Purux 'man'	purbɔ 'earlier'	Purbopurux	Ancestor, primogenitor
Jatra 'journey'	Xubha 'décor'	Xubhajatra	Procession

Table 14: *Compound Noun Compounds in Assamese (N+Adj : N)*

4.7.c. Compound Noun in Assamese with Adj +N : N

Head 1 (Adj)	Head 2 (Noun)	Compound Noun	Meaning
Lajuki 'shy'	Lɔta 'creeper'	Lajuki Lɔta	Memusa pudica

Losekelo (2019) in Kiswahili points out that the plant name kifauongo 'shame-plant (Memusa pudica)' obtains its name in Kiswahili from its physiological behavior. It is a sensitive plant that responds to touch and other stimulation by rapidly closing its leaves and drooping.

Compound Noun formation with Verb + Verb: Noun structure

This is also a compounding process of nouns available in both the languages. The noun formed by this process is rare and so not many examples could be found.

4.8.a. Compound Noun in Boro with V + V: Noun

Head 1 (Verb)	Head 2 (Verb)	Compound Noun	Meaning
Zan 'to distance'	Buŋ 'to say or tell'	Zanbuŋ	Cell phone

Table 15: *Compound Noun Compounds in Boro (V+V : N)*

4.8.b. Compound Noun in Assamese with V + V: Noun

Head 1 (Verb)	Head 2 (Verb)	Compound Noun	Meaning
Ana (to bring)	Kata (cut)	Anakata	traditional hand-woven cotton fabric specially woven by Assamese women

Table 16: *Compound Noun Compounds in Assamese (V+V : N)*

4.9. Compound Noun formation with Verb + Noun: Noun

4.9.a. Compound Noun in Boro with V+N : N

Head 1 (V)	Head 2 (N)	Compound Word	Meaning
khan (to comb)-	zɔŋ (spear)	khanzɔŋ	Comb
maɔ (to work)	thaŋkhi (aim)	maɔthaŋkhi	Action plan
san (to imagine)	simaŋ (dream)	Sansimaŋ	Fantasy
za (to happen)	thai (fruit)	Zathai	Incident
dankhɔ (to extract)	thai (fruit)	dankhɔthai	Extraction
muzu (to spit)	duɯi (water)	muzuduɯi	Spit

Table 17: *Compound Noun in Boro with V+N : N*

4.9.b. Compound Noun in Assamese with V + N : N

Head 1 (Verb)	Head 2 (Noun)	Compound Noun	Meaning
Ura 'fly'	jahaz 'ship'	Urajahaz	Aeroplane
Sura: ŋ (smuggle)	Sikari 'hunter'	Surangsikari	Smuggler

Table 18: *Compound Noun in Assamese with V+N : N*

Compound Adjectives

Compound adjectives are very rare in both languages. A few encountered examples of compound adjectives in both the languages are given below.

Head 1	Head 2	CompoundAdjective	Meaning	Language
dan(month)	sraŋ (bright)	dansraŋ	Moonlit	Boro
aatmɔ 'self'	sɔnma:n 'esteem'	aatmɔ sɔnma:n	Pride	Assamese
shrɔjɔn 'create'	xɔkti 'strength'	shrijɔni xɔkti	Prolific	Assamese
xirɔ (head)	bisxed (to cut)	xirɔ bisxed	decapitate, behead	Assamese

Table 19: *Compound Adjectives in Boro & Assamese*

dan(month)-adj.sraŋ (light)> adj.dansraŋ (moonlit).

The word dansraŋ is constructed linking to the nature of the Moon. When the Moon appears in the form of a full moon the light lit by it becomes bright and shiny. As moonlit is attributed by the cyclical nature (of a period of one month) of the Moon the two words dan (month) and sraŋ(bright) have been compounded to denote moonlit.

Other variations of compound formations seen in Assamese and Bodo

4.8.a (i) Other variations of compound formations seen in Bodo

Head 1 (Noun)	Head 2 (Verb)	Compound Word	Meaning
dau (fowl)	duɯi (to lay egg	daudɯi	Egg
ha (soil)	zɯu (be high)	hazɯ	Mountain
ha (soil)	sib (to sweep)	hasib	Broom
si (cloth)	ba (to take on back)	sibba	an wooden or bamboo frame used for keeping cloths
san (sun)	duŋ (to heat)	Sanduŋ	Sunlight
bhum (earth)	su (to measure)	bhumsu	Geometry

Table 20: *Variations of compound formations seen in Bodo*

4.8.a (ii) Compound Nouns formed with Three Free Morpheme in Bodo

Head 1	Head 2	Head 3	Compound Noun	Meaning
Noun	Noun	Noun	Noun	
bar (wind)	duɯi (water)	Sikhla (damsel or young girl)	bardɯisikhla	goddess of storm and wind
Noun	Verb	Noun	Noun	

ɔkha (rain)	ha (to rain)	dɯi (water)	ɔkhahadɯi	rain water/ rain drop
Pronoun	Verb	Noun	Noun	
bi (he/she)	hɯ (to give)	thai (fruit)	bihɯthai	Contribution

Table 21: *Compound Nouns formed with Three Free Morpheme in Bodo*

4.8.a (iii) Compound Verb

Verb + Verb: Verb

za (to eat) +phɯi (to come)> v. zaphɯi (come and eat)

Verb + Verb + Verb: Verb

maɔ (to do) + zɯb (to finish) + phɯi (to come)> v. maɔzɯbphɯi (come and do completely)

Verb +Verb +Verb +Verb: Verb

Za (to eat) + laŋ (to take away) + zɯb (to finish) + phɯi (to come) > v. zalaŋzɯbphɯi (come and eat completely) etc.

In this compounding structure the meaning of the compounded verbs are expressed collectively to denote different extended meanings of the newly constructed word.

4.8.b. Other variations of compound formations seen in Assamese

Head 1	Head 2	Compound	Category
pɔd 'post'	unnɔti 'progress'	pɔdunnɔti 'promotion	N+V = V
mrityu 'death'	dɔndɔ 'punishment'	mrityu dɔndɔ 'death sentence'	N+N = V
subhɔ 'good'	aagɔmɔn 'arrive'	subhagɔmɔn 'good arrival'	Adj+V : Adj
ɔgni 'fire'	ba:n 'arrow'	ɔgniba:n	N+N :Adj
mukh 'face'	sɔndrɔ 'moon'	mukhsɔndrɔ	Moonfaced

Table 22: *Variations of compound formations seen in Assamese*

Hybrid Compound

Bodo language also has accepted the process of hybridization to some extent to expand its vocabulary.

Mathigaldab is a hybrid word formed by the combination of two words – mathi and galdab. Mati is an Assamese word denoting soil and galdab is Bodo word which means flat. Mathigaldab is that variety of edible leafy vegetables found rooted on soil spreading its flat leaves.

Narasiŋa is another hybrid word that composed of nara (Bd. straw) and siŋa (Ass. to reap). Narasiŋa is the seasonal festival related to harvest of crop that performed in Bodo society during harvest. After this festival farmers do not have any work for harvesting again during the season. Because, at the time of narasiŋa, all crops are harvested successfully.

Besides the above mentioned hybrid words the language uses some hybrid words that contain two different words from two different languages. Neither of the components belong to Bodo language. A few specimens are-

Sanskrit. rasa (liquid) + Bangla. gɔlla (round shaped dish) > rasagɔlla (one kind of Indian sweet dish)

Hindi. lathi (stick) + Eng. charge (formal accusation)> n. lathicharge (an action taken by police personnel to scatter demonstrators by using lathi (stick) especially beating them)

Conclusion

Studies on the properties of compounds (see Benczes 2006; Libben & Jarema 2006; Lieber & Štekauer 2009; Scalise et al. 2009; Kageyama 2010; Scalise & Vogel 2010; Moyna 2011; Štekauer et al. 2012; Appah 2017) substantiate that compounds exist in the world's languages as a mechanism of word-formation.

The analysis has thrown light on the similar processes of compound word formation used in Boro and Assamese language. It is noteworthy that neighboring Assamese language has an etymological relation to this word bar-dɯi-sikhla for its word bɔr-dɔi-sila whose meaning is storm in its language. Bar-dɯi-sikhla is popularly known and believed to be a goddess having a power to create storm in the sky. According to a traditional belief of Bodo, when she comes on earth, comes in the form of a storm flying her black hair in the wind in a devastating mood.

Furthermore, this research paper explores the role of word formation in shaping language acquisition, cognitive processes, and communication strategies. By examining how speakers navigate the vast repertoire of words available to them through various word formation processes, this study aims to deepen our understanding of how language structures thought, perception, and social interactions.

In conclusion, the investigation of word formation in languages serves as a gateway to unraveling the intricate tapestry of human communication. Through a comprehensive analysis of the mechanisms and principles that underpin word creation, this research paper aims to contribute to the broader field of linguistics by illuminating the fundamental processes that drive language diversity and innovation.

References

Bauer, Laurie. 2001, Morphological Productivity. Cambridge, UK: Cambridge University Press.

Toman, J. 1992 "Compounding" In International Encyclopedia of Linguistics. ed. Brown, W. Oxford University Press, New York.

Daimai, K. (2019). Word formation in contemporary Liangmai: A morphological study. Himalayan Linguistics, 18(2).

Bhattacharya, P.C. 2007. A Descriptive Analysis of the Boro Language, University publication Department, Gauhati University.

Snyder, William (in press) Compound word formation. Jeffrey Lidz, William Snyder, and Joseph Pater (eds.) The Oxford Handbook of Developmental Linguistics. Oxford: Oxford University Press.

Baro, Madhu Ram, 2007. The Historical Development of Boro Language, N.L. Publications, ABR Road, Panbazar,Guwahati-781001

Shameem, Tanvir. 2016. August. Word-Formation Process.

Basumatary, Jahnabee. Gender in Bodo and Assamese Language: A Comparative Study Jahnabee Basumatary

Chainary, S.P. 2014 Tibeto-Burman Languages of North East India, Mittal Publications, New Delhi, India.

Crystal, David. 2008, A Dictionary of Linguistics and Phonetics, Blackwell publishing 350 Main Street, Malden, MA 02148-5020, USA

Endle, Rev. S, 1997. The Kacharis, Low Price Publications, B-2. Vardhaman Palace, Nimri Commercial Centre. Ashok Vihar, Phase-IV, Delhi-110052
Thakur, D. 1997 Linguistics simplified Morphology, Bharati Bhawan, Thakurbari Road, Kadamkuan, Patna- 800003
Madhu Ram Baro, Gwjwo Raokhanthi (Hajo, Kamrup, Priyadini Brahma, 1996) 110.

Jadavpur Journal of Languages and Linguistics
ISSN: 2581-494X

Derivational Adverbs by adding Locative Case Marker '*-aw*' in Bodo

Gojen Swargiary[1] *& Biswajit Brahma*
Cotton University, Assam

ARTICLE INFO

Article history:
Received 12/09/2025
Accepted 06/10/2025

Keywords:
Bodo language,
adverb,
locative case marker,
Bodo derivational adverb

ABSTRACT

This article examines how derivational adverbs are formed in Bodo, a Tibeto-Burman language, by adding the locative case marker "-aw". These adverbs fall into two categories: adverbial of place and time, derived from nouns, pronouns, and adjectives. For instance, nɔwaw (in the house), bejaw (here), pʰuŋaw (in the morning), and sanaw (in the day). With the personal and interrogative pronoun sɯr, the genitive -ni precedes -aw, as in sɯr-ni-aw (whom). This process illustrates the productive use of case marking in Bodo grammar.

Introduction

The Bodo language is an important language spoken in Assam and nearby areas of India. It belongs to the Tibeto-Burman group of the Sino-Tibetan family and is one of the oldest languages in the region. Bodo was originally an oral language used to pass down stories and traditions. In the late 20th century, the Devanagari script was adopted for writing, and it has been used ever since. Before that, in the late 19th century, some foreign scholars and native writers also used the Roman, Bangla, and Assamese scripts for writing. Today, Bodo is one of the 22 scheduled languages of India and has official status in the Bodoland Territorial Region. It is taught in schools and used in radio, television, and newspapers. However, it faces challenges from other languages like Assamese and English. Organizations like the Bodo Sahitya Sabha work hard to preserve and promote the language.

Review of Literature

In the history of the Bodo language, before the missionaries came to Assam, there are no written records on grammar or literature in the language. As evidence, the book Outline Grammar of the Kachari (Boro) Language as Spoken in District Darrang, Assam (1884) by Rev. Sidney Endleis considered the first grammar of the Bodo Language. This book provides an early introduction to the grammar and structure of Bodo, including the adverbs. Kamal Kumar Brahma categorized adverbs into four types in his book Gwnang Raokhanthi (1972) but no more description of the formation of adverbs. Later, Dr. Pramod Chandra Bhattacharya, in A Descriptive Analysis of the Boro Language (1977), describes adverb formation in the Bodo language and mentions the use of the locative case to form derived adverbs. Besides these early works, further linguistic contributions of native linguists and scholars also described adverbs and derived adverb formation. For instance, according to Swarna Prabha Chainary (Boro Raokhanthi, 2006), adverbs

[1] Corresponding Author Email ID: ginswargiary@gmail.com

in Bodo can be categorized into three types: adverbs of time, manner, and place. In general, adverbs in Bodo are typically placed before verbs. On the other hand, Aleendra Brahma (Modern Bodo Grammar, 2013) classified adverbs into two types from the point of view of how they are formed: basic/primary and derived. Basic or primary adverbs are three types. Adverbs that exist as independent words are called basic adverbs rather than being derived from other forms. These are adverbials of time, place, and manner. Derived adverbs are formed by adding the derivational affix {-ɯi} to other parts of speech or to itself. Daimalu Brahma also categorizes adverbs formed by suffixation into two types in his book Word Formation Processes in Bodo: Affixation (2017): adverbs of manner and adverbs of location. Adverbs of manner are derived from adjective bases by adding the derivational suffix /-ɯi/. Adverbs of location are formed by adding the morpheme /-ao/ to a noun, pronoun, or adjective.

However, there is still scope for studying how locative case markers function in forming derivational adverbs in Bodo. The locative case marker -aw can be combined with other parts of speech to form derivational adverbs.

Methodology

In this research, a descriptive-analytical method has been applied to analyze the transformation of nouns, pronouns, and adjectives into adverbs by adding case markers in Bodo and examining how these adverbs are used in Bodo sentences.

The data has been collected from primary and secondary sources.

Objectives

The main objectives of the study are as follows:

i. To analyze the formation of derivational adverbs by adding locative case markers.
ii. To examine the usage of derivational adverbs in sentences.
iii. To identify special characteristics of locative case markers.

Discussion

The word ‘adverb’ comes from the Latin word ‘adverbium.’ This word is made up of two parts: ‘ad-’ (which means ‘to’ or ‘addition’) and ‘verbum’ (which means ‘word’ or ‘verb’). True to its etymological roots, an adverb is a part of speech that modifies the meaning of a verb, adjective, or another adverb. It also conveys information about the time, place, and manner of an action described by a verb. According to the Oxford Learner’s Dictionary, an adverb is “a word that adds more information about place, time, manner, cause, or degree to a verb, an adjective, a phrase, or another adverb.”[2] The Cambridge Dictionary defines an adverb as “a word that describes or gives more information about a verb, adjective, adverb, or phrase.”[3]

The formation of adverbs is an element of morphology, which is the study of the smallest grammatical units in a language and their role in forming words, including inflectional and derivational structures. Ten cases are found in Bodo: nominative, accusative, instrumental, dative, ablative, genitive, locative, possessive, allative, and illative. All the cases have at least one case marker. However, the locative case marker is -aw.

[2] https://www.oxfordlearnersdictionaries.com/definition/american_english/adverb
[3] https://dictionary.cambridge.org/dictionary/english/adverb

Derived adverbs by adding -aw

An adverb is a word that modifies or describes a verb, an adjective, or another adverb. It provides additional information about how, when, where, to what extent, or in what manner something happens. The case markers are the inflectional nominal suffixes, and they can be attached to nouns, pronouns, numeral classifiers, adjectives, and adverbs in the sentence. But sometimes, some case markers attached to nouns, pronouns, and adjectives take roles as derivational suffixes and form a new word. Many Bodo adverbs are formed by adding the derivational suffix to other parts of speech or adverbs themselves. As is well known, various ways of forming adverbs exist in the Bodo language. This paper describes how adverbs in Bodo are formed by adding the locative case marker -aw. Two types of adverbials — adverbial of place and adverbial of time—are derived using this marker. These adverbs take the form of Base + Case Marker, where the base can be a noun, pronoun, or adjective. The primary case marker involved in this process is the locative marker -aw.

In the Bodo language, the locative case marker -aw has three allomorphs: -aw, -jaw, and -waw. The form -aw occurs with nouns ending in a consonant or the vowel /ɯ/, whereas -jaw is used with nouns ending in a front vowel or central vowel, and -waw with those ending in a back vowel (/ɔ/and /u/).

Adverbial of Place

The adverb which indicates the place of action or state is called the adverbial of place.[4] These kinds of adverbs are derived from nouns, pronouns, and adjectives. Examples of such derived Bodo adverbs are provided below.

<u>From noun</u>

The adverbs of place are derived from nouns using -aw (Base + case markers). For example-
nɔ ‘house’+ -aw > adv. nɔwaw ‘in the house’
lama ‘road’+-aw > adv. lamajaw ‘on the road’
ɔkʰraŋ ‘sky’+-aw > adv. ɔkʰraŋaw ‘in the sky’
hazɯ ‘hill’+-aw > adv. hazɯaw ‘on the hill’
baizɯ ‘out’+-aw > adv. baizɯaw ‘outside’ etc.
In the sentence:
In the following example, the adverbial of place /nɔwaw/ (in the house) modifies or describes the verb /kʰam-dɯŋ/ (burning). It also answers the question: where is the fire burning? /nɔwaw/ (in the house).

(1)
nɔwaw ɔr kʰam-dɯŋ
in the house fire burn-CONT
(It’s on fire in the house)

In Bodo, these adverbs appear in sentences where the adverbial words /lamajaw/, /ɔkʰraŋaw/, and /baizɯaw/ denote places or locations of a verb. For example,

(2)
a. lamajaw gari salai

[4] Brahma, Aleendra. *Modern Bodo Grammar.* N.L publications,2013 p 126

on the road car drive
(Drive the car on the road)

b. ɔkʰraŋaw zɯmɯi bir-dɯŋ
in the sky cloud fly-CONT
(Clouding in the sky)

c. baizɯaw da-tʰaŋ
outside Neg. go
(Don't go outside)

From pronoun

"A term used in the grammatical classification of words, referring to the closed set of items which can be used to substitute for a noun phrase (or single noun)."[5] Seven types of pronouns can be found in Bodo: personal pronouns, demonstrative pronouns, interrogative pronouns, indefinite pronouns, reflexive pronouns, reciprocal pronouns, and relative pronouns. Among these pronouns, demonstrative pronouns and interrogative pronouns directly take locative case markers except for 'sɯr' (interrogative pronoun) in the formation of adverbs in Bodo. The analysis of using locative case markers (-aw) with pronoun is given below:

Be 'this'+-aw > adv. bejaw 'here'
bɯi 'that'+-aw > adv. bɯijaw 'there'
ma 'what'+-aw > adv. majaw 'where'
bɔbe >mabe >ɔbe 'which'+-aw > adv. bɔbejaw > mabejaw > ɔbejaw 'where'
In the sentence:
Examples of the use of /bejaw/, /bɯijaw/, and /majaw/ in the sentences are given below.

(3)

a. bejaw sɯr dɔŋ?
here who have
(who is here?)

b. bɯijaw pʰɔraisapʰɯr dɔŋ
there students have
(Students are there)

c. bizab-a majaw dɔŋ?
book-NOM where have
(where is the book?)

The locative case is used with all the personal pronouns, as pronouns are also used as locations in sentences. But locative case markers, when used with personal pronouns and interrogative pronouns 'sɯr'(who) are preceded by the genitive case marker -ni with the locative marker -aw, as shown below. But most of the time, {-nijaw} (GEN+LOC, that is, -ni + -aw) is used as -naw.

aŋ 'I'+-ni + -aw> aŋnijaw / aŋnaw 'to me'
zɯŋ 'we'+-ni + -aw> zɯŋnijaw / zɯŋnaw/ zɯŋhanaw 'to us'
nɯŋ 'you'+-ni + -aw> nɯŋnijaw/ nɯŋnaw 'to you'
nɯŋsɯr 'you'+-ni + -aw>nɯŋsɯrnijaw/ nɯŋsɯraw 'to you'
bi 's/he'+-ni + -aw> binijaw/binaw 'to her/him'
bisɯr 'they'+-ni + -aw> bisɯrnijaw/bisɯrnaw 'to them'

[5] Crystal, David. *A Dictionary of Linguistics and Phonetics*. Blackwell publishing (6th ed 2008) p 391

nɯŋ-thaŋ ‘you’+-ni + -aw> nɯŋthaŋnijaw/ nɯŋthaŋnaw ‘to you’
nɯŋ-thaŋ-mɯn ‘you’+ -ni + -aw> nɯŋthaŋmɯnnijaw/nɯŋthaŋmɯnnaw ‘to you’
bi-thaŋ ‘s/he’ +-ni + -aw> bithaŋnijaw/ bithaŋnaw ‘to her/him’
bi-thaŋ-mɯn ‘they’ + -ni + -aw> bithaŋmɯnnijaw/ bithaŋmɯnnaw ‘to them’
sɯr ‘who’+ -ni + -aw> sɯrniaw/sɯrnaw ‘whom’
In the sentences:
Examples of the use of /aŋniaw/aŋnaw /, / zɯŋnijaw/zɯŋnaw/zɯŋhanaw /, and / sɯrniaw/sɯrnaw / in the sentences are given below.

(4)

a. aŋniaw / aŋnaw raŋ dɔŋ
to me money have
(I have money)

b. zɯŋnijaw / zɯŋnaw / zɯŋhanaw mablaba berai-nw pʰɯi
to us sometimes walk-SUB come
(Come to our home sometime for roaming)

c. nɯŋ sɯrniaw/sɯrnaw sɯŋ-dɯŋ?
you whom ask-PFV
(whom do you ask?)

Sometimes preceded by -naw, -ha is also used. -ha is the possessive case marker that expresses possession in the sentence. But, in this sentence, it is used as a location.

<u>From adjective</u>

sa ‘upon’+-aw > adv. sajaw ‘over’
siŋ ‘in’+-aw > adv. siŋaw ‘inside/under’
gɯzɯo ‘high’+-aw > adv. gɯzɯowaw ‘above’
gahaj ‘low’+-aw > adv. gahajaw ‘below’
gezer ‘middle’+-aw > adv. gezeraw ‘in the middle’
sigaŋ ‘before’+-aw > adv. sigaŋaw ‘before’
un‘back’+-aw > adv. unaw ‘behind’
mɯkʰaŋ ‘front’+-aw > adv. mɯkʰaŋaw ‘in front of’
setʰi ‘first’ +-aw > adv. setʰijaw ‘in the first’
In the sentences:
Examples of using / sajaw /, / siŋaw /, and /gɯzɯowaw/ in the sentences are given below.

(5)

a. bipʰaŋ-ni sajaw dau bir-laŋ-dɯŋ
tree-GEN over bird fly-take way-CONT
(Birds are flying over the tree)

b. nɯŋ gelegra pʰɯtʰar-ni siŋaw pʰɯi
you playground-GEN inside come
(You come inside the playground)

c. sɯŋtʰibilai-ni gɯzɯowaw bilai nɔŋ hɯ
Question paper-GEN above page number give
(Put the page number above the question papers)

d. gahajaw hɯ-nai sɯŋlu-pʰɯr-kʰɯo mɯzaŋ-ɯi pʰɔrai
below give- question-PL- ACC good-SUFF read
(Read the questions carefully given below)

Adverbs of Time

The adverb which indicates the time of action or state is called adverbial of time[6]. Many adverbials of time are formed from nouns and end in -aw. The Basic rule is that -aw is added to the end of the noun:

From noun

pʰuŋ 'morning'+-aw > adv. pʰuŋaw 'in the morning'
pʰuŋbili 'early morning'+-aw > adv. pʰuŋbilijaw 'in the early morning'
san 'day'+-aw > adv. sanaw 'in the day'
sanzwupʰu 'noon' +-aw > adv. sanzwupʰuwaw 'at noon'
belasi 'afternoon'+-aw > adv. belasijaw 'in the afternoon'
mɯnabili 'evening'+-aw > adv. mɯnabilijaw 'in the evening'
hɔr 'night'+-aw > adv. hɔraw 'at night'

Sometimes come numeral with nouns to form derivational adverbs such as given below:
san 'day'- se 'one' +-aw > adv. sansejaw 'in a day'
san day'- nɯi 'two' +-aw > adv. sanɯijaw 'in the two days'
habtʰa 'week'-se 'one' +-aw > adv. habthaseaw 'in a week'
dan 'month'- se 'one' +-aw > adv. dansejaw 'in a month'
bɯsɯr 'year'-se 'one' +-aw > adv. bɯsɯrsejaw 'in a year' Etc.
In the sentences:
Examples of using / pʰuŋaw /, / sanaw /, and / hɔraw/mɯnajaw / in the sentences are given below.

(6)

a.	pʰuŋaw	tʰabainai-a	deha-ni	tʰakʰai	mɯzaŋ
	in the morning	walking-NOM	health-GEN	for	good
	(Doing a morning walk is good for health)				
b.	nɯŋ	aŋ-kʰɯo	lɯgɯ	homɯ-bla,	sanaw	pʰɯi
	you	I-ACC	friend	meet -SUB	in the day	come
	(If you meet me, come in the day)					
c.	dinɯi	hɔraw/mɯnajaw	ɔkha	hanɯ	hagɯu	
	today	at night	rain	can	may	
	(It may rain today at night)					

Conclusion

In Bodo, adding locative case markers is one of the processes used to form adverbs. Derivational adverbs are primarily formed from nouns, pronouns, and adjectives. This process results in two types of derivational adverbs: adverbial of place and adverbial of time. These adverbs are created by combining a base with a case marker. For adverbials of place, the base can be a noun, pronoun, or adjective, and they are formed by attaching locative case markers to these bases. Examples include baizɯ-aw> baizɯaw (outside), bɯi-aw> bɯijaw (there), and sa-aw> sajaw (over). Similarly, adverbials of time are derived by adding a locative case marker to a noun base, as seen in pʰuŋbili-aw >pʰuŋbilijaw (in the early morning) and hɔr-aw > hɔraw (at night).

[6] Brahma, Aleendra. *Modern Bodo Grammar.* N.L publications, 2013 p 125

However, when locative case markers are used with personal pronouns and the interrogative pronoun sɯr, the genitive case marker -ni precedes the locative marker -aw to form adverbials of place, as in aŋni-aw > aŋnijaw / aŋnaw (to me) and nɯŋni-aw > nɯŋnijaw / nɯŋnaw (to you). These derivational adverbs play a crucial role in Bodo grammar, enriching its expressive capability and demonstrating how case marking extends beyond nouns to influence adverbial structures.

References

Baro, Madhu Ram. (2007). The Historical Development of Boro Language (New ed.). Guwahati: N.L. Publications.
Basumatary, Guddu Prasad. (2014). A Descriptive Analysis on Bodo Suffix (unpublished thesis), Gauhati University, Guwahati. http://hdl.handle.net/10603/116204
Basumatary, Phukan Chandra. (2000). Boro Raoni Mohorkhanthi. Baganpara, B.T.A.D Assam: Neelima Prakashani.
Bhattacharya, Pramod Chandra. (1977). A Descriptive Analysis of the Boro Language.Guwahati: Gauhati University Press.
Brahma, Aleendra. (2013). Modern Bodo Grammar. Guwahati: N.L. Publications.
Brahma, Aleendra. (2020-21). "Case Systems in Bodo and Dimasa." GUINEIS JOURNAL, Pages: 232-255.
Brahma, Daimalu. (2017). Word Formation Processes in Bodo Affixation. Satmile, Near Gauhati University: Purbayon Publication.
Brahma, Kamal Kumar. (1972). Gwnang Raokhanthi. Panbazar, Guwahati: Bina Library.
Chainary, Swarna Prabha. (2015). Boro Raokhanthi (Fifth edition, Revised). Baganpara, Neelima Prakashani.
Chainary, Swarna Prabha and Basumatary, Phukan Chandra. (2023). Gwdan Boro Swdwb Bihung. Baganpara, Neelima Prakashani.
Crystal, David. (1992). A dictionary of linguistics and phonetics. Blackwell Publisher.
Hockett, Charles F. (1958). A course in Modern Linguistics. New Delhi: Mohan Primlani Oxford & IBH Publishing Co.
Mahela, R. & Sinha, S. (2018). "Case Markers in Bodo and Sanzari Bodo: A Comparative Investigation." International Journal of Linguistics.

Abbreviation

ACC Accusative case
CONT Continuous tense
GEN Genitive case
HAB Habitual aspect
LOC Locative case
NEG Negative
NOM Nominative case
PL Plural
PFV Perfective aspect
SUB Subordinating Conjunction
SUFF Suffix

Jadavpur Journal of Languages and Linguistics
ISSN: 2581-494X

Affixations in Kokborok

Priya Debbarma[1]
Tripura University, Tripura

ARTICLE INFO

Article history:
Received 15/09/2025
Accepted 06/10/2025

Keywords:
Kokborok,
Tibeto-Burman,
word Formations,
affixations,
prefixes,
suffixes

ABSTRACT

Kokborok is a Tibeto-Burman language spoken mainly in the north-eastern part of India, Tripura. Words are formed through various processes like inflections, derivations, compounding, reduplications etc. This paper mainly focuses on creation of new words through affixations. No infixes and circumfixes are found in Kokborok. Both prefixes and suffixes play a significant role in word formations processes through affixations. Thus, this paper focusses on detailed discussion of affixations in Kokborok.

Introduction

Kokborok belongs to the Tibeto-Burman sub family, linguistically. This language is spoken mainly in Tripura, and some parts of Assam and Mizoram, the northeastern state of India and other countries like Bangladesh and Mayanmar. The term kokborok is a compound of two words, kok means ‘language’ and borok means ‘people’. Generally, kokborok refers to the language of people. The speakers of Kokborok refer to themselves as Tiprasa or Tripuri. The word ti comes from tui means water, pra comes from bupra means confluence and sa from the word sa means son. As per the census report of 2011 the total populations of Kokborok speakers are 10,11,294.

The language, Kokborok, was put in different branches or sections by different scholars. G.A Grierson in his book Linguistic Survey of India (1903), classified Kokborok under Bodo-Naga group. Robert Shafer (1966-1974) made a classification of Sino-Tibetan languages dividing Tibeto-Burman languages into four main groups; these include Bodic, Burmic, Karenic and Baric. According to him Kokborok belongs to the Western Units of the Barish Section within the Baric sub division of Sino Tibetan. Burling (2003) also classified Kokborok under the sub group of Bodo languages which is a group of Bodo-Koch part of Bodo-Konyak-Jingphaw group of Tibeto-Burman languages. Kokborok is an SOV word order highly agglutinating in nature i.e. head final language. This language shares a typological feature with Bodo, Garo and Dimassa.

Research Objectives

The main aim of this particular study is to find out the affixes that are employed in this language. It also focuses on analysing and discussing the functions and occurrence of affixes in Kokborok.

[1] Corresponding Author Email ID: priyadebbarma361m@gmail.com

Methods of data collection

The data is collected basically from the primary source and some data is also collected from secondary sources gathered from books like journals, articles etc. For primary data, an interview has been made by visiting the west district of Tripura in the area of Khumulwng, Jangalia and Jampuijala. As a native speaker of Kokborok, the researcher also collected data by interacting among the informants and by observing how little information had been collected from the speakers.

Discussion

Affixation is a morphological process of forming new words with the help of affixes. The affixes are added to the root or the stem of the word. New words are formed with the help of prefixes and suffixes. In Kokborok word formed through suffixes is more productive than prefixes.

Prefixation

Prefix occurred at the beginning of a root or stem of a word. Kokborok has a nominal prefix and a verbal prefix. Nominal prefixes are the prefixes that occur only with nouns whereas verbal prefixes occur along with verbs.

Nominal prefixes

Nominal prefixes can be formed with the help of pronominal markers in Kokborok. Pronominal markers play a very significant role in the processes of Kokborok word formations. Some of the kinship terms like -ma, -pha, -cui, -cu, -ta form the pronominal prefixes with the help of 1st person, 2nd person, and 3rd person in Kokborok. The following examples are discussed below,

(1)

a	+	ma	>	ama	‘my mother’
a	+	pha	>	apha	‘my father’
a	+	cəi	>	acəi	‘my grandmother’
a	+	cu	>	acu	‘my grandfather’
a	+	ta	>	ata	‘my brother’
nə	+	ma	>	nəma	‘your mother’
nə	+	pha	>	nəpha	‘your father’
nə	+	cəi	>	nəcəi	‘your grandmother’
nə	+	cu	>	nəcu	‘your grandfather
nə	+	ta	>	nəta	‘your brother’
bu	+	ma	>	buma	‘his/her mother’
bu	+	pha	>	bupha	‘his/her father’
bu	+	cəi	>	bəcəi	‘his/her grandmother’
bu	+	cu	>	bucu	‘his/her grandfather’
bu	+	ta	>	buta	‘his/her brother’

It is clearly seen in the examples above that the pronominal prefixes are attached to the noun to create new words. And these pronominal prefixes can occur only with the kinship terms with the

help of 1st person, 2nd person, 3rd person to indicate the meaning of 'my', 'your' and 'his/her'. It is also observed that without pronominal prefixes such kinship stands only for mother, father, sister, grandfather, grandmother, brother.

Verbal Prefixes

In Kokborok, the verbal prefix is very limited. Verbal prefixes in Kokborok occur only with negative prefix -ta. For examples,

(2)

ayaŋ-o	ta-thani
there-LOC	NEG-go

'Don't go there.'

(3)

nəŋ	omo-no	ta-ca-di
You	this-ACC	NEG-eat-IMP

'You don't eat this!'

(4)

bui-ni		manəi	ta-taŋ-di
other-	GEN	thing	NEG-touch-IMP

'Don't touch other thing!'

It is observed from the above sentences that prefix ta- which means 'don't' is attached to the verb root along with the imperative marker -di.

Suffixation

Suffixation in Kokborok is very productive in nature. It occurs in a larger number than prefixes. Suffixes in Kokborok can be added to nominal root and verbal root. Suffixation also plays an important role in the processes of word formation in Kokborok.

Nominal suffixes

Nominal suffixes are those suffix that can be attached with noun and pronoun only. Nominal suffixes in Kokborok are attached with gender marker and plural marker in Kokborok. The examples are, Nominal suffixes for gender marker

The suffixes -la, -cəla, -Jəla, -sa denote male genders. As in the example,

(5)

Masculine	**English translation**
tok-la	cock
sui-la	male dog
takhum-cəla	drake
məsa-cəla	male tiger
sa-Jəla	son
musuk-Jəla	bull

From the example given above it is very clear that suffixes -la, -cəla, -□əla, -sa are employed in the language to denote male gender. They are used in different noun entities i.e., suffix -la is used for male sex of animals like birds, dogs etc., suffix -cəla is attached to animals like drake, male elephant, male tiger, and suffix -□əla is added to identify both human and animal gender.

The suffixes -Jək, -ma, -bərui denotes to female gender marker in Kokborok.

(6)

Feminine	**English Translation**
phərə□nai-Jək	female teacher
sa-Jək	daughter
pu-ma	mother goat
wak-ma	mother pig
tokthu-bərui	dove
amiŋ-bərui	female cat

It is clear that when suffixes -□ək, -ma, -bərui is added to the noun it denotes the meaning of female gender. The suffix -Jək is employed to both young and adult female gender markers, suffix -ma is added to indicate the meaning of mother goat, mother pig etc and when suffix -bərui is attached, refers to young female gender marker.

Nominal suffixes for plural marker

The suffixes -rog and -soŋ are marked for nominal suffixes in Kokborok. Examples are,

(7)

Singular	**Plural**	**Gloss**
borok	borok-rog	people
musuk	musuk-rog	cows
cəla	cəla-rog	boys

In Kokborok the plural suffix -rog can occur with both animate and inanimate objects as in the example given above.

The suffix -soŋ is used only for human. It cannot occur with animate and inanimate objects. As can be discussed below:

(8)

Singular	**Plural**	**Gloss**
bupha	bupha-soŋ	fathers (father and his friends)
bai	bai-soŋ	sisters
ram	ram-soŋ	Ram and his friend

When -soŋ is used with proper noun it expresses the concerned entity and its groups and -soŋ also expresses honorific noun when it is used with kinship terms.

Verbal suffixes

In Kokborok verbal suffixes are those that are attached to the verbal root. The nature of these suffixes are bound i.e. they cannot stand on their own. New words are derived by adding suffixes to the verb. Verbal suffixes in Kokborok are nominalising morpheme, adverbial marker, causative suffix, negative suffix and reciprocal suffix.

Nominalising morpheme

The nominalising suffixes in Kokborok are *-nai, -thai, -muŋ* and these are discussed in details below;

When the suffix -nai is added to the verb it acts as a noun. The nominaliser -nai occurs to the right of the verb. As in the example given below,

(9)	**(10)**
phərəŋ ‘teach’	sərəŋ ‘learn’
phərəŋ-nai	sərəŋ-nai

teach-NMLZ
‘Teacher.’
The suffix -nai as a nominaliser.

learn-NMLZ
‘Learner.’

The nominaliser -thai expresses the ‘place of the action’ performed by the verb. The examples are,

(11)
acai ‘born’
acai-thai
born-NMLZ
‘Place of birth.’
The suffix -thai as a nominaliser.

(12)
toŋ ‘stay’
toŋ-thai
stay-NMLZ
‘Place to stay / ‘lodging.’

(13)
sərəŋ ‘learn’
sərəŋ-thai
learn-NMLZ
‘Place to learn.’

When nominalising suffix -muŋ is added to verb it indicates the way of doing something. Examples are

(14)
ca ‘eats’
ca-muŋ
eat-NMLZ
‘Way of eating.’
The suffix -muŋ as a nominaliser

(15)
toŋ ‘stay’
toŋ-muŋ
stay-NMLZ
‘Way of staying.’

It can be said from the above example that the suffixes *-nai, -thai, -muŋ* are the nominalisers that are added to the verb to form deverbal noun in Kokborok.

Adverbial marker

In Kokborok the adverbial marker -khe is attached to the verb to denote the manner of the adverb. Examples are,

(16)

tərək-khe	>	‘slowly’
siriŋ-khe	>	‘quietly’
naithok-khe	>	‘beautifully’

In the example above when the suffix -khe is added to the verb, new words are formed denoting the manner of adverb in Kokborok.

Negative suffix

The suffix -ya, -gəlak function as a negative marker in Kokborok. This negative suffix is bound morpheme, when it is attached to the verbal root it denotes negations. Examples are,

(17)

phai+ya (come + NEG)	>	phaiya	'don't come'
səi+ya (write+ +NEG)	>	səiya	'don't write'
sok+ya (rotten +NEG)	>	sokya	'not rotten'
ca+ gəlak (eat + NEG)	>	'cagəlak	'may not eat'
thaŋ+ gəlak (go +NEG)	>	'thaŋgəlak'	'may not go'

It can be analysed from the above example that when the suffix -ya is added with the root verb it denotes the meaning of 'not'. Whereas the suffix -gəlak is added to the root verb to denote the meaning of 'may not' as in the example given above.

Reciprocal suffix

In Kokborok the reciprocal suffix -lai is added to the verb to derive new words. Examples are given below,

(18)

walai + lai	>	walailai	'quarrel with each other'
bulai + lai	>	bulailai	'fight with each other'
khitar + lai	>	khitarlai	'throw with each other'
cirik + lai	>	ciriklai	'shout with each other'

It is clearly seen that when the reciprocal suffix -lai is added to the verb it also indicates the meaning of plurality as given in the above examples.

Conclusion

It can be analysed from the above examples that new words are derived through the process of affixation in Kokborok. Kokborok has prefixes and suffixes. The nominal prefixes are formed with the help of pronominal markers in Kokborok. The language makes use of three pronominal prefixes i.e. 1st person, 2nd person and 3rd person and this pronominal prefix occurs only with kinship terms in Kokborok. The verbal prefixes in Kokborok occur only with the negative prefix -ta. Words formed through prefixes are very limited in Kokborok. While suffixes play a very significant role in the word formation processes in Kokborok. The nominal suffixes include plural marker, gender marker, and case marker. Verbal suffixes are nominalising morpheme, adverbial marker, causative suffix, negative suffix and reciprocal suffix in Kokborok.

References

Bauer, L, (2004). English Word-Formation. Cambridge University Press.

Debbarma, Binoy, (2002). Anglo- Kokborok-Bengali Dictionary; Agartala, Kokborok tei Hukumu Mission.

Haspelmath, Martin & Sims, D. Andrea, (2010). Understanding Morphology: 2nd edition published by Hodder Education and Hachette UK Company.

Ingty, Angela R.W, (2008). Garo Morphology: A Descriptive Analysis. Ph.D Dissertation, North Eastern Hill University. [Dept. of English]

Jacquesson, Francois, (2008). Kokborok Grammar [Agartala Dialect]. Agartala: Kokborok Tei Hukumu Mission.

Kattamba, F & J. Stonham, (2004). Morphology. 2nd Edn. New York, Palgrave Macmillan.

Karapurkar, Pushpa P, (1979). Kokborok Grammar. Mysore: Central Institute of Indian Languages.

Plago, Ingo, (2003). Word Formation in English. Cambridge, Cambridge University Press.

Singh, S.I, (2021). Word Formation in Manipuri. Indian Linguistics, Vol. 82 Nos 1-2, P. 91-105.

Abbreviations

ACC	Accusative
ADV	Adverbial
GEN	Genitive
IMP	Imperative
LOC	Locative
NMLZ	Nominaliser
PL	Plural
REC	Reciprocal

Jadavpur Journal of Languages and Linguistics
ISSN: 2581-494X

Case Syncretism in Assamese through Caha's Framework: A Synchronic Analysis

Dewashree Ojah[1] *& Arup Kumar Nath*
Tezpur University, Assam

ARTICLE INFO

Article history:
Received 12 /09/2025
Accepted 06/10/2025

Keywords:
Syncretism,
Assamese,
dative,
accusative,
instrumental,
locative,
allative

ABSTRACT

Syncretism is defined language-internally in terms of a subset of lexemes belonging to a particular part of speech having a form corresponding to more than one morphosyntactic feature value. In historical linguistics, it refers to the identity between two forms of the same lexeme. Many researchers have described Syncretism in their terms, yet there is limited exploration of Syncretism within a specific theoretical framework in Indian languages. (Das, 2020) mentioned the presence of case Syncretism in Assamese. Though a lot of work has been done on Assamese Linguistics, the area of syncretism has not been studied particularly in the language. Assamese is the easternmost Indo-Aryan language, spoken primarily in the Brahmaputra valley. The objective of this paper is to investigate the morphosyntactic structure of Syncretism in Assamese among various Case Markers. Here is an example of the -k marker, which is used to indicate both Accusative and Dative cases. e.g.,

1) xi tai-k pit-il- e
3sg.m 3sg.f-acc beat-pfv-3
"He has beaten her."

2) xi tai-k putola e-ta di-l- e
3sg.m 3sg.f-dat doll one-clf give-pst-3
"He gave her a teddy."

In sentence (1), tai (3sg.f) takes the accusative case -k as the patient of the sentence; however, in sentence (2), we can see that tai (3sg.f) takes the dative case -k as the recipient of the doll. As we can see, both the dative and accusative markers are originally distinct in terms of their syntactic function; however, the features are being syncretised here.

The paper depends on both primary and secondary sources. The paper attempts to contribute to the area of morphosyntax in Assamese. Moreover, it would include a discussion on different case markers by giving more emphasis on those markers that can report several functions. The paper mainly examines Case syncretism in Assamese by utilising (Caha, 2009) Universal Case Contiguity Principle and Panini's Karaka Theory. It explores non-accidental Case syncretism in Assamese, aligning with(Caha's,2009) argument that such syncretism occurs in neighbouring Case domains following a universally consistent order across languages.

Introduction

(Blake, 2001) defines case as a "System of marking dependent nouns for the type of relationship they bear to their heads". It is a grammatical category that specifies the precise function a word serves within a sentence. These inflectional markers can be applied to nouns, pronouns,

[1] Corresponding Author Email ID: diojah98@gmail.com

adjectives, numerals, and, in certain languages, adverbs, indicating their respective roles in the sentence structure. Assamese, an Indo-Aryan language spoken in North-East India, features a complex case system. It serves as the official language of Assam and acts as a common language among the diverse linguistic communities in the region.

In Assamese, subjects exhibit varied case forms, typically contingent on the characteristics of the subject. Due to its complex case system, there is a lack of morphological distinction among the different case markers in Assamese. As a result, the same case marker is used to mean the originally distinct inflectional forms, either through the merger of the forms or through the merger of their underlying function. This phenomenon is known as Syncretism(Pott, 1833) According to (Caha, 2009) Case syncretism is a situation in which "two separate Cases share the same form." Nevertheless, it is often challenging to determine whether the syncretism between Case markers is a random occurrence or happens systematically and deliberately. According to (Das, 2020) Assamese shows syncretism between (a) the accusative and dative, (b) between possessive, experiencer, and genitive, and (c) between instrumental and comitative.

The primary goal of this paper is to investigate various empirical and conceptual aspects of case syncretism in Assamese. This paper focuses on identifying explicit case markers, emphasising thematic roles, and proposing grammatical relationships.

Research Methodology

This study is not merely descriptive; it is an investigative exploration; therefore, it necessitates minute observation of the language to fulfil the objectives. The study includes various aspects of morphology, syntax, semantics, and discourse. The methodology is primarily qualitative. The research is dependent on both primary and secondary sources. This work combines theoretical analysis with a data-driven perspective. The study employs an inductive methodology to provide an empirically grounded analysis of case marking in Assamese. To achieve the objectives, various methods such as observation, direct and indirect personal interviews, and fieldwork have been conducted. Secondary sources for data collection have included descriptive grammar books, newspaper articles, and digital archives of Assamese texts.

Aims and Objectives of the Paper

1. To explore different empirical and conceptual aspects associated with case typology in Assamese
2. To study case markers, highlight their thematic roles and suggest grammatical connections.
3. To investigate the morphosyntactic structure of Syncretism in Assamese Case Markers.
4. To observe whether Caha's (2009) Universal Case Contiguity principles are applicable in Assamese or not.

Theoretical Framework

In this paper, primarily two theories will be utilised: Panini's Karaka Theory from the 4th BC and Caha's Universal Case Contiguity from 2009.

Karaka Theory: Panini’s grammar ‘Astadhyayi’ serves as a primary modelling device with its fundamental principles of affixation, base, affix, deletion, modification, and intra and inter-category relations. In Paninian tradition, the concept of Karaka has been highly emphasised. There, the verbs function as a clause’s head, and each nominal dependent is assigned a karaka.

The relationship between 'bibhakti'(case) and 'karaka' (semantic role) in Panini's framework is complex. There is no direct one-to-one correspondence between the 'karakas' (semantic roles) and case markings. In this system, case markers are merely numbered instead of being directly linked to specific semantic meanings. Panini stated that Sanskrit features six 'karakas' and seven bibhaktis for nouns. Here is a table showing the correspondence between Panini's Bibhaktis and Western Case:

Bibhakti	**Case**
Prathama Bibhakti	Nominative Case
Dwitiya Bibhakti	Accusative Case
Tritiya Bibhakti	Instrumental Case
Chaturthi Bibhakti	Dative Case
Panchami Bibhakti	Allative Case
Sasthi Bibhakti	Genitive Case
Saptami Bibhakti	Locative Case

Table 1: *Correspondence between Panini's Bibhaktis and Western Case*

Panini believes that the definition of morphological case is influenced by several semantic factors. The Karakas, their definition, associated semantic role, and related inflectional markers in Assamese can be established by following traditional definitions. The table below shows the karakas, their definition, related bibhktis, and related semantic roles:

Karaka	**Related Semantic Role/Case**	**Related Bibhaktis**	**Definitions**	**Related Inflectional Marker in Assamese**
Karta	Agent	Prathama	The doer or the performer	-a,-ɛ
Karma	Patient	Dwitiya	The entity that is directly affected	k/ ɔk
Karan	Instrument	Tritiya	The means or instrument	-re/-di/-dwara
Sampradana	Goal	Chaturthi	The entity that receives something	-loi
Apadan	Ablative	Panchami	The point from which something departs or is separated	-pɔra
Sambandh	Genitive	Sasthi	Possession of a close association	-r/-ɔr
Adhikarana	Locative	Saptami	The location	-t/-ɔt

Table 2: *The karakas, their definition, related bibhktis, and related semantic roles*

Universal Case Contiguity Principle: Caha's proposal on Case syncretism describes it as the phenomenon where "two distinct Cases have the same form" (Caha, 2009) However, determining whether the syncretism between Case markers is accidental or systematic is not always straightforward. Caha argues for a systematic approach to Case syncretism, claiming that it involves universal mechanisms and is not merely a result of phonological similarities on the surface syncretism or a matter of accidental homophony. Syncretism in languages can sometimes be mistaken for accidental homophony. Caha provides an example from Greek to illustrate this point. In Greek, the syncretism between nominative and accusative, as well as between accusative and genitive cases, is not merely coincidental. Instead, it occurs in a systematic and principled manner.

(Caha, 2009) discusses a specific type of Case syncretism that is distinct from phonological processes or accidental homophony. He introduces the concept of systematic, non-accidental syncretism, which is governed by a particular constraint known as the Universal Contiguity Principle. According to this principle, only contiguous cases in a universal Case order can exhibit syncretism.

Universal Case Contiguity:
a) Non-accidental case syncretism targets contiguous regions in a sequence invariant across languages.
b) The Case sequence: nominative - accusative - genitive - dative - instrumental - comitative (Caha, 2009)

This principle anticipates that syncretism will be observed between nominative and accusative, accusative and genitive, genitive and dative, and so forth. However, it does not predict syncretism between non-contiguous regions, such as nominative and genitive by excluding accusative, or dative and nominative, while excluding Cases in between.

Existing Literature

In traditional accounts, Assamese consists of seven cases. According to traditional Assamese grammars, influenced by Sanskrit grammars, depict the case system as a nominative-accusative one. (Bora, 1960) identifies six types of karakas and seven types of bibhakti in Assamese. According to (Bora, 1960) Korta functions as the doer and the grammatical subject of the sentence. Prathama bibhakti is assigned to Korta karaka, indicating the grammatical object of the sentence. Karma typically represents the grammatical object that is affected by the verb's action. Karan corresponds to the instrumental case, while Sampradan corresponds to the dative case. Apadan refers to the origin, root, or source of things. Adhikaran corresponds to the locative case, denoting place. (Bora, 1960)mentions the term sambandha as 'genitive' but does not consider it a karaka, as it lacks direct relation with the verb. In contrast, Golokchandra (Goswami, 1982)outlines seven cases with their morphological markers, introducing the concept of 'absolute case' in Assamese. The absolute case lacks overt markers and is utilised for both subjects of a specific class of intransitive verbs and direct objects of transitive verbs. Additionally, it is employed for objects of passive verbs. (Goswami, 1982)mentions seven grammatical cases in Assamese: absolutive, nominative, accusative, genitive, instrumental, dative, locative. But (Nath & Barbora, 2016) mentions about 'ergative case'. As with many other languages, there is syncretism in Assamese case markers. In his Ph.D. (Das, 2020) mentioned the presence of Syncretism within Assamese case markers.

(Richa Srishti, 2011) Focuses on the key issue of Case Syncretism, which is one of the biggest challenges to the annotation of corpora in Indian languages. (Haspelmath, 2023) established a systematic set of terms in a cross-linguistic context as an umbrella term for syncretism, polysemy, cogrammification, coexpression, syllexification, syngrammification, and synexpression. Another paper, "Marking the Unmarked: Exceptional Patterns of Syncretism in English and Hindi" by Paul Arsenault, addresses a particular approach to morphosyntactic features. In this approach, the 'least marked' feature in a morpho-syntactic dimension (e.g., person, number, case, etc.) is equated with the 'least specified' feature structure. As a result, the paper establishes that the unmarked features are always underspecified about marked features and lack unique feature representations. (Das, 2020) mentioned about the presence of Syncretism within Assamese case markers, found among. (a) accusative and dative cases, (b) possessive, experiencer, and genitive cases, as well as (c) instrumental and comitative cases in the Assamese language.

As with many other languages, there is syncretism in Assamese case markers. But this area of Syncretism has not been studied yet within a theoretical framework.

Syncretism in Accusative Case and Dative Case k/-ɔk

In the Paninian framework, karma is associated with the grammatical object in a sentence. This indicates that in Assamese, the direct object of a transitive verb is identified as being designated with the dwitiya bihakti, commonly known as the 'accusative case.'

On the other hand, in the context of Assamese traditional grammar, the dative case corresponds to chaturthi bibhakti, which is linked to the semantic role of "handing over." In ancient Greek, the dative case is used to indicate the indirect object with verbs expressing the act of giving. In Karaka Theory, the term sampradan signifies giving something to someone permanently, signifying a transfer of authority from the donor to the recipient. To illustrate, activities such as lending money, paying taxes, or giving clothes to a washerman do not fall under the category of sampradan. On the contrary, offering goods to a priest or giving something to a beggar exemplify instances of sampradan. In Assamese traditional grammar, there are two markers for the dative case: "loi" (allomorph "ɔloi") and "-k" (allomorph "-ɔk"). For example:

1) xi tai-k pit-il- e
3sg.m 3sg.F-acc beat-pst-3
"He has beaten her."

2) moi tai-k pit-il-u
1sg tai-acc beat-pst-1
"I have beaten her."

3) xihot-e bhoniyek-ɔk mat-il-e
3pl-erg sister-acc call-pst-3
"They called their sister."

4) xi tai-k putola e-ta di-l- e
3sg.m 3sg.f-dat doll one-clf give-pst-3
"He gave her a teddy."

5) moi Ram-ɔk bɔh-ib-loi soki-k^{h}ɔn di-l-u

1sg Ram-dat sit-fut-dat chair-clf give-pst-1
"I gave the chair to Ram to sit on."

6) xihot-e kukur-tu-k kha-bo-loi di-l-e
3pl-erg dog-clf-dat eat-nf-dat give-pst-3
"They gave food to the dog."

7) xihot-e kukur-ɔk kha-bo-loi di-l-e
3pl-erg dog-dat eat-nf-dat give-pst-3
"They let the dog eat."

In the examples above, the markers -k and ɔk are used as allomorphs. The -k marker is applied to vowel-ending pronouns such as -tai, while for consonant-ending nouns like Ram and Kukur, the allomorph -ɔk is used instead. The sentences (1), (2) and (3) take the accusative case -k/ɔk as the patient of the sentence; however, in the sentences (4), (5), (6) and (7) we can see that tai takes the dative case -k/ ɔk as the recipient of something. As we can see, both the dative and accusative markers are originally distinct in terms of their syntactic function; however, the features are being syncretised here.

Syncretism in Ergative and Locative Case Markers - e

Agentivity plays a key role in many discussions of ergativity. In Assamese, human nouns, personal names, and kinship terms can be marked ergative. Regarding pronouns, singular pronouns and 1st person plural pronouns are not marked with ergative, whereas 2nd and 3rd person plural pronouns are overtly marked. Honorific pronouns are also marked with the ergative case. In the Assamese traditional account, -ə, -e, and -i are stated as the markers for Prathama bibhakati. For example:

8) ram-e b^hat kha-l-e/ pela-l-e/ rand^h-i-l-e / t^{h}o-l-e
3sg- erg rice eat-pst-3/throw-pst-3/ cook-nf-pst-3/keep-pst-3
"Ram ate/threw/cooked/kept the rice"

9) xihɔt-e hutel-ɔt b^hat kha-l-e
3pl-erg hotel-loc rice eat-pst-3
"They had dinner at the hotel."

10) tumaluk-e b^hat kha-l-a
2sg-ergrice eat-pst-2
"You ate rice."

In the above examples, we can see that the -e marker is used to indicate ergativity in Assamese. Here, the subjects receive the theta role of an agent of the action and take the ergative case -e. Plural subject pronouns, namely the second person tʊmɑlʊk 'you-plu' and the third person xihɔt/tɑhɑt 'they' also take ergative case as shown in the above example.

11) boroxun-e tiai/ d^{h}ui nis-e
3sg-erg drenched take-3
"We are drenched in rain."

12) otipaat bororxun-e banpani kore
excessive rain-erg flood does

"Excessive rainfall causes flooding."

In the above examples (11) and (12), we can see that even inanimate objects like Boroxun "rain" also receive the Ergative case as it takes the agent theta role.

In the Assamese locative case expresses the notion of location. The locative case in Assamese corresponds to Saptami bibhakti and is associated with adhikaran karaka. Mostly, the noun is inflected with case markers and marks the place where the action is carried out. In the example below, we can see that the sense of location is marked in the reduplicated noun with the help of the - e marker. e.g.,

13) dal-e dal-e kɔpau pʰul pʰul-is-e
branch-loc branch-loc orchid flower bloom-pfv-3
"Orchids are blooming on the branches."

14) tumi din-e din- e xokot hoi goi as-a
2sg day-locday-locfat become go exis-2
"You are becoming fat day by day."

From the above sentences, it can be seen that the -e marker is used to denote location. But at the same time a question may arise if reduplication has any role to play here, as in both sentences it can be seen that the subjects are being reduplicated. To clear the doubt, we can take some other examples as well:

15) lora-tu din- e barh-is- e
3sg.clf day-locgrow-pfv-3
"The boy is growing every day."

16) gos-jupa dal-e pat- e xuka-is- e
tree-clf brance-loc leaf-loc dry-pfv-3
"The tree has dried from branches to leaves."

Even in the absence of reduplication, a sense of location can still be observed in the above sentences. In colloquial Assamese, we use sentences like (15) and (16) to give emphasis on a particular situation. However, in a broader context, it indicates the sense of location.

Thus, in the above sentences, we have seen that the same marker-e is being used to denote both Ergativity Case and Locative Case in Assamese.

Syncretism in Allative Case and Dative Case Marker -loi in Assamese

In Assamese, the allative case is used to indicate movement towards a specific location or destination. It signifies that an action is directed towards a certain location. For example, the sentences below (17), (18) and (19) indicate some movements towards a spatial goal, which is why they denote the allative case.

In Paninian grammar, dative corresponds to chaturthi bibhakti and is associated with the semantic role sampradan, to hand over, purpose, cause, goal, direction, etc. (Rabha Hakacham, 2015). The dative case includes three-place verbs in Assamese, such as di 'give', agbarha 'offer', dekhuwa 'show', etc. In Assamese, -loi/ɔloi and -k/ɔk denote the Dative Case.

17) moi Guwahati-loi goi as-u/o
1sg 3 sg-all go aux-1
"I am going to Guwahati."

18) gʰɔrɔ-loi ɔha alɔhi-k ɔnadɔr kɔr-ibɔ na-pai

home-all come.nom guest-acc disrespect do-fut neg-get
"You should not disrespect a guest who comes to your house."

19) hozabo-loi ɔna p^hul-bur xotez hole b^hal
decorate-all bring flower-PL fresh be. pst.3(cond)good
"It's good if the flowers brought for decoration are fresh."

20) tumi sita-loi ki anisa?
2sg 3-dat what bring
"What have you brought for Sita?"

21) teu oha-loi-ke moi bat sam
3sg come-dat-until 1sg wait see
"I will wait until he comes."

In (17), (18) and (19), the action suggests movement towards a spatial destination, hence requiring the Allative Case. However, in (20) and (21), there is a distinct indication of giving or offering something to someone, necessitating the Dative Case.

Syncretism in Locative Case and the Nominalised-t marker in Assamese

Again, in the following sentences, we observe the marker "-t" undergoes syncretism after displaying distinct behaviour in the two instances provided below. In (22) and (23), there is a clear indication of the Locative Case.

22) mur ghɔr Roha-t
1gen home Roha-loc
"I live in Roha."

23) boroxun-ɔt titi tar jor ho-is-e
rain-loc drench 3sg.m cold being-asp-3
"He caught a cold after getting drenched in the rain."

24) xi khuwa-t bhal
3sg eat-nmlz good
"He is good at eating."

25) tumi oha-t mur bhal lag-il
2sg come-nmlz 1sg good feel-pst
"I'm glad that you came."

But in sentences (24) and (25), it becomes apparent that "-t" attaches after the nominalised form "/oha/". However, in this context, the "-t" marker cannot be interpreted as a locative case. Instead, its inclusion suggests a conditional sense apart from Nominalizer.

Syncretism in Instrumental Case/ Allative Case

The instrumental case serves to indicate the noun representing the instrument or tool employed by the agent to accomplish an action. The instrumental case corresponds to Tritiya Bibhakti in Karaka Theory, which is associated with karan karaka. -re/-ere) and -di/- edi) are the inflectional

instrumental markers in Assamese. In Assamese, there are two methods of expressing the instrumental case. One approach is used in active voice, employing suffixes like -re/-ere, while the other involves the use of postpositions such as -edi/-di, dwɑɹɑ, hɔtuɑi, zɔɹiɔte, and others.
The sense of instrument is conveyed also by the instrumental -e followed by the conjunctive participle -di "giving"; e.g.hat-edi (kakati, 1935)Sometimes, Instrumental case is also marked by -di, which is called Tritiya Bibhakti in Paninian Grammar (Sharma, 2022)

Allative Case indicates movement 'through something'. mɑz-edi 'through the middle' gɑʊ- kʰɔn-edi 'through the village' bʰitɔr-edi 'from inside' are some examples of the allative case in Assamese.

26) ei nɔdi-r buku-edi bɔhu pani bagɔr-i go-l
this river-gen chest-all many water flow-nfgo-pst
"A lot of water flowed through this river."

27) xi pʰulɔni-r mɑz-edi ɑh-is-e
3sg garden-gen middle-all come-pfv-3
"He is coming through the forest."

28) tai haat-edi bʰat kʰa-bɔ par- e
3sg hand- inst rice eat-fut-3
"She can eat rice with her hands."

29) kutʰar- edi gos-dal kat-il- e
axe-inst tree-clf cut-pst-3
"(He) cut the tree with an axe."

In the example above, the same marker -edi is functioning in two different roles, in the first two sentences it is being used as an Allative case marker, and in the last sentences the -edi marker is used as an Instrumental case marker.

Syncretism in Instrumental Case and Comitative Case

The comitative case expresses the meaning of 'along with' or 'accompanied by'. In Assamese, 'xoite', 'logot', etc., are some of the common expressions to denote comitative meaning. The example below shows the comitative case in Assamese:

30) tumi ki-here bʰat kʰa-l-a
2sg what-com rice eat-pst-2
"With what did you write?"

31) tumi ki-here ghɔr-tu lip-i-l-a
2sg what-com house-clf plaster-nf-pst-2
"With what did you plaster the house?"

In the example above, the ere marker denotes the comitative case in Assamese. But as we can see, a consonant /h/ is inserted, and it becomes here. According to (Das,2020), ki 'what' is inflected by the -ere morpheme to express the comitative meaning. Before -ere is added to ki, there is an instance of consonant epenthesis (Blevins, 2008) where /h/ is inserted between /i/ and /e/.
However, the same-ere marker is the allomorph of -re marker and is used to indicate the instrumental case as well. For example,

32) sima-i pensil-ere sobi ak-i-l-e
3sg.erg pencil-com picture draw-nf-pst-3
"Sima drew a picture with a pencil."

33) tai hat-ere phul-pah siŋ-il-e
3sg hand-com flower-clf pluck-pst-3
"She plucked the flower with her hands."

Syncretism in Genitive Possessive and Experiencer

The genitive case corresponds to sasthi bibhakti in Assamese. Chomsky considers the genitive as an inherent case. The genitive marker in Assamese is the suffix -r. Its allomorph -ɔr is phonologically conditioned. Like other languages, Assamese also has identified a large class of meanings associated with experiencer construction (Nath & Barbora, 2016)These include- State of sickness, State of consciousness, Physical state, Emotional state, State of things in one's situation, Inalienable Possession, etc.

For example:

34) mur jor ho-is e [State of sickness]
1sg-gen fever is-pfv-1
"I have a fever"

35) rina-r laj lag-is- ɛ [State of consciousness]
3sg-gen shame feel-pfv-3
"Rina feels shy."

36) mur bhagor lag-is-e [Physical state]
1sg.gen tired feel-pfv-1
"I am feeling tired."

37) ram-ɔr a: nandɔ lag-is-e [Emotional state]
3sg-gen happy feel-pfv-2
"Ram is feeling happy."

38) a:mar du-k^{h}ɔn ha:t as-e [Inalienable possession]
1pl.gen two-clf hand exis-3
"We have two hands."

39) boby-r du.ta kɔlɔm as-ɛ [alienable possession]
3sg-gen two-clf pen exix-3
"Boby has two pens."

In the examples above, it can be observed that they are marked by the genitive marker r/ɔr, yet they express different states of experiencer constructions in Assamese.

Findings and Conclusion

Assamese has a very complex morphological case system. This paper has attempted to demonstrate some of the complexities and systematic properties of case markers in Assamese. From the examples above, it has been observed that, in addition to the definite markers mentioned in traditional Assamese grammar, we use some other markers as well. Following the theta rules, these additional markers play a crucial role in conveying the meaning of a specific situation. Thus, it cannot be asserted that certain inflectional markers are exclusively used for a specific grammatical case. The table below shows the additional inflectional markers that can be used to denote Grammatical Cases in Assamese:

Grammatical Case	**Theta Role**	**Related Inflectional Marker in Assamese**	**Added Inflectional Marker to Denote Meaning**
Ergative Case	Agent	-a/-ɛ	
Accusative Case	Patient	-k/-ɔk	
Instrumental Case	Instrument	-re/-ere ,-di/-ɛdi, dwara	
Dative Case	Goal	-loi	-k/ -ɔk
Allative Case	Allative	-pɔra	-di/-ɛdi, -loi
Genitive Case	Source	-r/ɔr	
Locative Case	Location	-t/-ɔt	-ɛ
Comitative Case	Association	logot. xoite	-re/-ere

Table 3: *Case Syncretism and additional grammatical case makers in Assamese*

It has been observed that Syncretism occurs very often among the case markers, where a single morpheme denotes different grammatical cases, such as

a. Ergative Case and Locative Case
b. Accusative Case and Dative Case
c. Instrumental Case and Comitative Case
d. Instrumental Case and Allative Case
e. Dative Case and Allative Case
f. Locative Case and the Nominalizer
g. Genitive Syncretism in Possessive and Experiencer.

Going by (Caha, 2009)Universal Case Contiguity principle, it is said that non-accidental case syncretism targets contiguous regions in a sequence invariant across languages. But in Assamese, we do not observe case syncretism occurring in a contiguous region, nor does it follow that particular sequence prescribed by (Caha, 2009). Additionally, it has also been mentioned that case sequence typically follows a definite order: nominative-accusative –genitive – dative –instrumental – comitative (Caha, 2009). However, in Assamese, Accusative Cases are syncretised with Dative Cases, leaving Instrumental Case in between. Similarly, the Instrumental Case and Allative Case are syncretised, leaving the Dative Case aside. The Ergative Case and Locative Case are syncretised, leaving out all the cases in between.

Only the Dative Case is syncretised with its adjacent case, the Allative Case. Even so, it cannot be concluded that Assamese is following (Caha, 2009)Universal Case Contiguity principle.

References

Blake, B. J. (2001). Case (2nd ed). Cambridge University Press.
Bora, S. (1960). Bahal Vyakaran (17th edition). Bani Mandir.

Caha, P. (2009). The Nanosyntax of Case. University of Tromso.
Das, P. (2020). Analysis of Assamese Morphology for a Rule-Based Assamese Spell Checker. Tezpur University.
Goswami, G. (1982). Structure of Assamese (1st edition). Dept. of Publication, Gauhati University.
Haspelmath, M. (2023). Coexpression and synexpression patterns across languages: Comparative concepts and possible explanations. Frontiers in Psychology, 14, 1236853. https://doi.org/10.3389/fpsyg.2023.1236853
Hjelmslev, Louis. (1935). La catégorie des cas: Étude de grammaire générale. Universitetsforlaget, 7.
Jakobson, R. (1936). Beitrag zur allgemeinen Kasuslehre. Gesamtbedeutungen: Der Russische Kasus.
kakati, banikanta. (1935). Assamese formation and development. Calcutta University.
Nath, D., & Barbora, M. (2016). The Agentive Ergativity in Assamese. Bhabani Offset Pvt. Ltd.
Pott, A. (1833). Etymologische Forschungen auf dem Gebiet der indo-germanischen Sprachen. Meyerschen Hofbuchhandlung.
Rabha Hakacham, U. (2015). Asomia Ruptottor Moulik Bisar. Assam Publishing Company. Guwahati.
Richa Srishti, S. M. (2011). Case Syncretism in Urdu-Hindi:A Challenge for NLP. CIIL Mysore.
Sharma, A. (2022). Bhasabijnan Aru Asomiya Bhasa (2nd edition). Bandhav.

Abbreviations

1	first person
2	second person
3	third person
ABL	ablative
ACC	accusative
ADJ	adjective
ALL	allative
AGR	agreement
AUX	auxiliary
CAUS	causative
CLF	classifier
COND	conditional
CONG	conjunctive participle
COP	copula
DAT	dative
ERG	ergative
F	feminine
FUT	future
GEN	genitive
IMP1	imperative honorific
IMP	imperative
INCL	inclusive
INF	infinitive
INS	instrumental
LOC	locative
NEG	negative
NMLZ	nominalizer
NOM	nominative
PASS	passive
PFV	perfective
PL	plural
POSS	possessive
PFV	perfective
PL	plural
PROG	progressive
PST	past
SG	singular

Jadavpur Journal of Languages and Linguistics
ISSN: 2581-494X

Multiple functions of *-k(a)*[1] morpheme in early Assamese

Gitanjali Bez[2]
Gauhati University, Assam

ARTICLE INFO

Article history:
Received 15 /09/2025
Accepted 06/10/2025

Keywords:
Genitive,
possessed nouns,
nominalized verbs,
Allative,
Dative subjects

ABSTRACT

-k morpheme is a very versatile morpheme which is used in different constructions and serves multiple functions in early Assamese. This paper discusses where it occurs and what kind of function is carried by it in each construction. The paper also presents a comparison of early Assamese and modern Assamese in regard to the uses of this morpheme. This comparison reveals that some of the dialects of modern Assamese are more conservative while some others are more progressive regarding the use of this morpheme.

Introduction

Assamese belongs to the eastern group of the Indo-Aryan language family and is spoken by a majority of people living in the state of Assam located in North-East India. It has a rich literary tradition which has started from the end of the 13th century by the writings of a few court poets. Prahrad Carita was the first text which has been claimed to be the earliest literary specimen of Assamese language (Kakati 1941). However, some features of Assamese language have shown up before 13th century in the text of Caryapada[3] and Sri Krishna Kirtan. The language of both the texts represents a mixture of dialects which has been assumed to be some standard Magadhan Apabhransa and "which in later times developed into distinctive Bengali and Assamese languages (Kakati 1941: 10)."

The early literatures of Assamese have been divided into two sections, considering the writings of Baishnavite poet Srimanta Sankardeva as a landmark, grouping the literatures composed prior to Sankardeva as 'Prāk Sankari Jugar Sāhitya' (the literatures of pre-Sankardeva period) and the literatures composed during his time as 'Sankari Jugar Sāhitya' (The literatures of Sankardeva period). The form of these literatures is poetic in nature. The renowned poet Mādhav Kandali, who rendered the Vālmiki's Rāmāyana into Assamese language, belongs to the pre-Sankardeva period. Kandali's Saptakānda Rāmāyana is considered as the earliest translation of Vālmiki's Rāmāyan into the regional languages of the Indo-Aryan family.

To discuss the multiple functions of -k morpheme, the writings from both pre-Sankardeva and Sankardeva periods are considered. The data for this study primarily come from three

[1] The vowel demarcated by () is used as a feature of poetic text. All the early texts are composed as verses. The syllable final 'a' vowel is pronounced to adjust the rhyming pattern of a verse when it is read/recited.

[2] Corresponding Author Email ID: gitanjalibez@gauhati.ac.in

[3] Caryapada is a collection of texts that are assumed to be written within the period of 7th century-13th century (śastri (Bangabda 1323)). The texts were collected by Mahāmahopādhāy Haraprasād śastri, a renowned Bengali scholar, from Nepal. The language of these texts is assumed to be Apabhransa bhāsā.

sources. The first source includes two computer readable corpora stored in CQPweb (https://cqpweb.lancs.ac.uk/) – (1) the corpus of Mādhav Kandali's Rāmāyana, and (2) Emille_CIIL Assamese Corpus. The former of these two corpora includes 266 written texts (204,615 word tokens and 22,232 word types) whereas the latter one includes 1,191 texts which have 3,005,337 word tokens and 223,333 word types. The second source is the old texts which are available in a non-computer-readable form or in a non-digital format. The third source is the old texts available in a digital format. It is noted here that the examples cited from old texts are represented in Roman transliteration. Contemporary examples are represented in IPA.
The organization of the paper is as follows. Section 1 presents a very brief background of Assamese language that is relevant for developing an understanding to the discussion provided throughout the paper. Section 2 deals with the constructions where -k occurs and with the functions that -k serves in each construction. A brief comparison of -k marked structure in early Assamese with modern Assamese is presented in Section 3. Section 4 summarizes the paper.

Multiple functions of -k in different constructions of early Assamese
-k is found to occur with an ordinary noun (NP), a possessed noun, a nominalized verb, and with a verb in early Assamese. In each occurrence, -k serves different kinds of functions which are dealt with one by one in the subsequent sub-sections.

-k with an ordinary noun
-k is used as a genitive morpheme with an ordinary noun in early Assamese, as shown in the following examples.

(1) {chānda-ka bāndha}
pleasure-gen bondage
'the bondage of pleasure.'
(Texts of Charyāpada: Charyā 1)

(2) cāh-ā cāh-ā āla baṛāẏi {yamunā-ka tīr-e}.
search-2fam search-2fam address word grandmother Yamuna-gen bank-loc
'Hey Grandmother, search (him) at the bank of Yamuna river.'
(Sri Krishna kirtan)

(3) {rāwaṇa-ka citta-ka} na-pār-i buj-ibā-ka
Rawan-gen mind-dat neg-can-imp understand-nmlz-gen
'It is not understandable in the mind of Rawana.'
(Mādhav Kandali's Rāmāyana: 4Kiskindhā21_CQPweb)

In the above examples, -k marks the dependent of the NPs demarcated by {} in which the position of the dependent noun and the head noun is filled by an ordinary noun.
Apart from this construction, -k marks an ordinary noun which is followed by a limited set of postpositions in early Assamese, as illustrated below.

(4) ēhisē {karma-ka lāgi} bha-il-i dēwarāja
this work-gen for become-pfv.2inf king of deity

‘(The sage Gautom said to Indra, the king of deity, with anger) You became the king of deity for this work.’

(Mādhav Kandali’s Rāmāyana: 1Adi38_CQPweb)

(5) kriṣṇa daraśan-e mo-ra mahā-lābh yā-ib-o{dwārakā-ka prati}
Krishna visit-loc 1sg-gen great-profit go-fut-1Dwaraka-gen towards
‘By visiting Krishna, I will be greatly profitable, (and thus I) will go towards Dwaraka’

(Kirtan_Srimanta Sankardeva)

In the structure demarcated by {} in (4) and (5), -k marks an ordinary noun followed by the postpositions lāgi and prati, which function as the head of the given structure. Here the given structure functions as an adverbial of reason.

Further, -k is used to mark an NP that functions as an object (only object of transitive verb, and direct and indirect object of ditransitive verbs) and a subject, as demonstrated below.

(6) matiẽ {thākura-ka} parinibit-ā
Minister king-dat check-3
‘The minister has checked the king.’

(Texts of Charyāpada: Charyā 12)

(7) {rāma-ka} {sītā-ka} di-l-ā yautuka apāra
Ram-dat Sita-dat give-pfv-3 dowry limitless or endless
‘All people gave endless dowry to prince Ram and princess Sita.’

(Mādhav Kandali’s Rāmāyana: 2Ayodhya1_CQPweb)

(8) a-prayas-e {daitya-ka} mār-il-ā dānt-e chir-i
neg-effort-loc monster-dat kill-pfv-2fam tooth-ins bite-sim
‘Without putting any effort, you killed the monster by biting with teeth.’

(Kirtan_Srimanta Sankardeva)

(9) kene citralekhā {kumara-ka} ni-l-e har-i
how Citralekha prince-dat bring-pfv-3 kidnap-cnv
‘How did Chitralekha kidnap the prince’

(Kumar Haraṇ kābya_Ananta Kandali)

(10) tēbēsē māg-ib-ōṁ bara yēhi { mō-ka} lāg-ē
only then ask-fut-1 boon what 1sg-datneed-3
‘Only then I will ask for whatever boon I need.’

(Mādhav Kandali’s Rāmāyana: 1Adi16_CQPweb)

In (6), (8), and (9), -k marks the only object NP of transitive verbs where in (7), it marks the indirect object NP of a ditransitive verb. The direct object in (7) remains unmarked. The -k marked NP in (10) functions as the subject. This kind of subject NP is known as dative subject construction in descriptive literature (Masica 1991). Many Indo-Aryan languages have this

construction (Masica 1991: 346-357). In Assamese (both early and modern Assamese), -k marked subject NP is warranted by a particular verb lāg 'need'.
-k also marks an NP which signals the directional meaning in early Assamese, as given in (10) and (11).

(10) niti niti dadhi bik-e {mathurā-k'} jā-e.
each day milk sale-3 Mathura-all go-seq
(He) goes to Mathura and sales milk each day.
(Sri Krishna kirtan)

(11) {Ayōdhyā-ka} ās-il-ā anēka mahāṛṣi
Ayodhya-all come-pfv-3 many great sage
'Many great sages came to Ayodhya.'
(Mādhav Kandali's Rāmāyana: 1Adi20_CQPweb)

(12) {Ayōdhyā-ka} ās-il-ā sabhāryye ranga-man-e
Ayodhya-all come-pfv-3 with.wife delightful-mind-advl
'(He) came to Ayodhya delightfully with his wife.'
(Kirtan_Srimanta Sankardeva)

In (10), (11), and (12), -k marks the NPs demarcated by {} and contributes the directional meaning to the relevant NP. The -k marked NP functions here as an adverbial of direction and the morpheme itself is used as an allative.

-k with a possessed noun
-k is added to a particular kind of nominal heads that denote kinship relationship in a possessive NP construction to index third person possessors. The example (13) illustrates it.

{bhāẏ-ēk-aka} mār-i kēnē rājya na-lawa-ẏa
younger brother-3poss-dat kill-seqwhy kingdom neg-take-3
'(Rawana asked Sita) why does Ram not kill his younger brother (Bharata) and take his kingdom?
(Mādhav Kandali's Rāmāyana: 3Aranya13_CQPweb)

The syllable initial vowel ē in this morpheme occurs due to phonological factors. A word that ends with a few vowels such as ẏ[4], u, diphthongs oi, ou and consonants warrants this sound to occur with -k. The occurrence of this -k is controlled by the third person possessors in early Assamese.

-k with a nominalized verb
-k occurs with a nominalized verb marked by -ibā, and serves different functions, as in ((13)-(17)). -k marked verb occurs in two constructions: (i) without a postposition ((13), (14), (15), (16)), and (ii) with a postposition ((17)). -k is used as a genitive morpheme in this type of construction.

(13) putra samē bas-il-anta {[yajña kar-ibā]-ka}

[4] It is noted here that it is the graphemic representation of the sound *i*. It is not the semi-vowel *y*.

son together with sit-pfv-3hon sacrificial ceremony do-nmlz-gen
'(The sage Bibhanda) sat with his son to do the sacrificial ceremony.'
(Mādhav Kandali's Rāmāyana:1Adi20_CQPweb)

(14) mānusa niuj-il-a {[mār-ibā]-ka' tāe}
man appoint-pfv-3 kill-nmlz-gen 3.acc
'Man has been appointed to kill him.'
(Sri Krishna Kirtan)

(15) {[bhakata-ka rākh-ibā]-ka} ās-il-ā prabandh-e
devotee-dat keep-nmlz-gencome-pfv-2fam diligent-advl
'(You) came diligently to keep your devotee safe.'
(Kirtan_Sri Sri Sankardev)

(16) di-l-õ hante bara {[jambu-dwipa jin-ibā]-ka}
give-pfv-1 cft boon Jambu-island defeat-nmlz-gen
'(I) would have given the boon of defeating Jambu island.'
(Ādicarit_Sri Sri Mādhabdev)

(17) {[tapa bhaṅga kar-ibā]-ka lāgi} cal-i ga-il-a.
meditation break do-nmlz-gen for move-cnv go-pfv-3
'(Fairy) went to break the meditation (need to check the meaning of tapa) of the sage Mandakarni (Māndakarni).'
(Mādhav Kandali's Rāmāyana: 3Aranya4_CQPweb)

All the NP demarcated by {} above, irrespective of whether they are followed by a postposition or not, serve the adverbial functions of purpose.

-k with a verb

-k is added to a finite verb to serve two functions: (1) to index third person subjects, and (ii) to express a request and a wish, as shown below.

(18) hanumant-ē {kar-il-ēka} laṅkā-ka dagadha
Hanumanta-erg do-pfv-3 Langka-dat burn
'Hanumanta burnt Langka.'
(Mādhav Kandali's Rāmāyana: 6Lanka54_CQPweb)

(19) ārśībbāda {kar-iẏō-ka} cal-i yā-ōṁ bana
blessing do-2fam-impv mive-cnv go-1 forest
'Give me your blessing. (I) am going to the forest.'
(Mādhav Kandali's Rāmāyana: Ay: Ch. 8; V. 16)

(20) .rāma-ra {kalyāṇa ha-uka } {mar-oka} rāwaṇa.
Ram-gen happiness be-opt kill-optRawana
'(I) wish that Ram be happy and Rawana be killed.'
(Mādhav Kandali's Rāmāyana: Laṅ: Ch. 48; V. 196)

In (18), -k forms a constituent with -ē, which together is used to index third person subjects. It is noted that this morpheme is still preserved in modern Assamese with a slight phonological change across the dialects to serve the same function (See § 3). The standard dialect has lost the final -k gradually and it preserves the initial vowel to index third person subjects, where some other dialects (Kāmrupiyā dialect, Barpetiā dialect) preserve -k with a change of -ē vowel to -ā (-a) for serving the same function.

A comparison of -k in early Assamese with the -k in modern Assamese

-k with an ordinary noun

The uses and functions of -k are confined in modern Assamese relative to early Assamese. The genitive use of this morpheme is preserved only in some dialects, such as in Kāmrupiyā dialect, Barpetiā dialect, in modern Assamese and that also in a particular type of constructions where the genitive marked noun is followed by a postposition similar to the examples (4) and (5) in early Assamese presented above (see § 2.1).

(18) oi, bappa oi, ei zia-to imman mɔta. ek
excl father excl this girl-clf too much man one

kʰut-oɹ {biskut-ok legi} gɔtai likira-hopa-ɹ lɔg-ɔt
piece-gen biscuit-gen for all children-clf.pl-gen company-loc

dɔn kɔɹ-ba ḍɔ-ss-i.
fight do-infv start-pf-3

'Oh my gosh, father, oh my gosh, this girl is too much mannish. She has started to fight with all the children for a piece of biscuit.'

(https://www.facebook.com/groups/nalbeirakothabotora/permalink/4101670156550485)[5]

Just because the word, which takes -k, ends with a consonant sound, -o is inserted before -k in this dialect for the ease of pronunciation. Furthermore, the early Assamese lāgi became legi in this dialect by the change of the vowel ā to e.

The place of -k in marking the relationship between a dependent noun and a head noun is taken in modern Assamese by another morpheme -ɹ, which was simultaneously used with -k in early Assamese. Gradually -k has started not to be used in this function and -ɹ becomes the only morpheme used for it, as shown below.

(19) ma-k-ɔɹ[6] gʰɔɹ-ɔt
mother-3poss-gen house-loc
'at mother's house'

(Self-elicited)

As shown above, the genitive use of -k was almost lost in modern Assamese apart from its presence in some complex postpositional constructions (see Example (18) from Nalbāriyā dialect, and (4) and (5) from early Assamese) in some dialects of it. In contrast, it is invariably used for the dative/accusative functions to mark (i) a particular kind of subject NPs which is

[5] This example is collected from a sub-dialect called Nalbariyā dialects of Kāmrupiyā dialect.

[6] -ɔ is inserted before -ɹ for the reason that the word where -ɹ is added ends with a consonant sound. -ɔɹ is an allomorph of the genitive -ɹ in modern Assamese.

warranted by the verb lag ‘need/want’ (the subject of needing) and which is known as dative subject constructions in Indo-Aryan literatures (Masica 1991), and (ii) some kinds of objects of a transitive and a ditransitive verb, as illustrated below.

(20) {mʊ-k} taka lag-e.
1sg-datmoney need-3
‘I need money.’

(Self-elicited)

(21) {sowali-zoni-k} maɪ-ibɔ kʰuz-is-a-neki?
girl-clf.HUM.f-acc kill-infv want-pf-2fam-ques
‘Do you want to kill the girl.’

(Emille-CIIL Assamese Corpus: nmm148a_CQPweb)

(22) stesɔn-ɔt sɔndɹɔkumaɪ-e {dibja-k} bibah bissed-ɔɪ
station-loc Candrakumar-erg Dibya-dat marriage divorce-gen
kagɔz-kʰɔn di-l-e
paper-clf.FLAT give-pfv-3

(Emille-CIIL Assamese Corpus: amr14_CQPweb)

The only object of the transitive verb in (21) and the indirect object of the ditransitive verb in (22) are marked by -k. The choice of using -k morpheme to mark an object in a transitive and ditransitive construction is triggered by some semantic properties such as animacy, specificity of an object NP along with the feature of whether a particular object NP is formed by a proper noun referring to an animate referent or an inanimate referent. Out of these three, the former two along with the proper nouns with an inanimate reference are loosely applied for determining the presence of -k, whether the proper noun with an animate reference necessarily requires -k to mark the relevant object NP. For example, any animate object NP or any specific object NP can be optionally marked by -k where an object NP formed by an animate proper noun obligatorily requires the presence of it, as illustrated below.

(23) { pʰɔɹasi sʊali-Ø}bia kɔɹ-a-i ɦi ameɹika-ɹ
France girl marriage do-tr-seq 3inf.m.sg Amerika-gen

nagoɹikɔttɔ gɹɔhɔn kɔɹ-il-e
citizenship accept do-pfv-3
‘After getting marriage with a France girl, he acquired the citizenship of America.’

(Emille-CIIL Assamese Corpus: amm26_CQPweb)

(24) {bɹansɔn-e sʊali-k} ekʊ bɛa bebohar nɔ-kɔɹ-e
Branson-erg girl-acc any bad behaviour neg-do-3
‘Branson does not behave badly with any girl.’

(Emille-CIIL Assamese Corpus: amm07_CQPweb)

(25) {sʊali-zon-i-Ø} mʊ-loi bia di-a
girl-clf.HUM-f 1sg-all marriage give-2fam
‘Give your daughter to me for marriage.’

(Emille-CIIL Assamese Corpus: ash41_CQPweb)

(26) {sʊali-zon-i-k} maɹ-ibɔ kʰuz-is-a neki?
girl-clf.HUM-f-acc kill-infv want-pf-2fam ques?
'Do you want to kill the girl.'
(Emille-CIIL Assamese Corpus: nmm148a_CQPweb)

(27) {ɔɦɔm-ɔk} umaɦɔŋkɔɹ-e bʰal pa-is-il
Assam-acc Umasangkar-erg good get-pf-pfv
'Umasangkar loved Assam.'
(Emille-CIIL Assamese Corpus: nmm123a_CQPweb)

(28) auɹaŋzeb-e ɹazput ɹɔza ɹamɦiŋhɔ-k pɔtʰa-l-e
Aurangzeb-ergRajput king Ramsingha-dat send-pfv-3

{ɔɦɔm-Ø} maɹ-ibɔ-ɹ kaɹɔne
Assam kill-nmlz-gen for
Aurongzeb sent the Rajput king Ramsingha for defeating Assam.'
(Emille-CIIL Assamese Corpus: atb05_CQPweb)

(29) ɹamsiŋhɔ-ɹ lɔg-ɔt ɔɦɔiŋkʰɔ ɦoinɔ ah-is-e
Ramsingha-gen accompany-loc many soldier come-pf-3

{ɔɦɔm-Ø} dʰɔŋɦɔ kɔɹ-ibɔ-loi
Assam destroy do-nmlz-all
'Many soldiers have come with Ramsingha to destroy Assam.'
(Emille-CIIL Assamese Corpus: atb05_CQPweb)

(30) tumi {ɹam-ɔk} dekʰ-is-il-a-neki?
2fam Ram-acc see-pf-pfv-2fam-ques
'Have you seen Ram?'
(Emille-CIIL Assamese Corpus: nmm94_CQPweb)

(31) *tumi {ɹam-Ø} dekʰ-is-il-a-neki?
2fam Ram-acc see-pf-pfv-2fam-ques
'Have you seen Ram?'
(Self-elicited)

The object NPs demarcated by {} in (23) and (24) are animate, non-specific. However, they are optionally marked by -k in (23) and (24). Similarly, the object NP demarcated by {} in (25) and (26) is animate, specific. However, the one in (25) is unmarked, but the one in (26) is marked by -k. Similarly, the proper noun referring to a place is optionally marked by -k in (27)-(29). In contrast, the proper noun referring to a person cannot be optionally marked by -k, as shown in (30) and (31). Without being marked by -k, the object NP demarcated by {} in (31) is ungrammatical.

The allative use of -k is preserved only in some dialects such as Kāmrupiyā and Barpetiā dialects. Example (32) illustrates it.

(32) azi ɔlɔ-m ɔlɔ-m buli ḇabɔ-t-e bɛla-e
today go out-fut go out-fut sub think-nmlz-loc half of a day-add

gɛ-l. ɦazbeila kaj zai-e {bazaɪ-ok}.
go-pfv evening who go-3 market-all
'The half of the day is over by thinking going out today. Who does go to the market in the evening time?
(https://www.facebook.com/groups/nalbeirakothabotora/permalink/4024355477615287)

This use of -k is completely lost in the standard dialect and replaced by a new allative -loi to serve the relevant function in the dialect, as in (33)

(33) ei godʰuli-kʰɔn {bozaɪ-oloi} kʊn go-is-e?
this evening-clf market-all who go-pf-3
'Who has gone to the market in this evening time?'
(Self-elicited)

-k with a possessed noun

Similar to the early Assamese, -k is added to kin terms to index third person possessors in modern Assamese (see Bez (2019)), as in (34).

(34) ta-ɪ pas-ɔt ta-ɪ ma-k aɪu
dem.dist-gen back-loc 3sg.inf.dist.m-gen mother-3poss and

aita-k utʰ-i ah-is-il.
grandmother-3poss rise-cnv come-pf-pfv
'After then, his mother and grandmother stood up and came back.'
(Emille-CIIL Assamese Corpus: nmm147_CQPweb)

Here, the kin terms ma and aita take -k to index the third person possessor tar.

-k with a nominalized verb

With nominalized verbs, the use of -k is lost entirely in the standard dialect of modern Assamese, but is still preserved in the dialects of Kāmrupiyā and Barpetiā, as demonstrated below.

(35) pɹɔtʰɔm niɦa-i apa-i ag-oɹ kɔtʰa-ɹ {pɹɔtiɦoḍ
first night-loc boy-erg front-gen subject-gen revenge

lɔ-ba-k legi} pʰɔndi koɹ-i kɔtʰa-tu
take-nmlz-gen for plan do-sim topic-clf

uilei-s-i.
raise-pfv-3
'The boy has raised the matter strategically at the first night only (he does not want to wait for the other day) to take the revenge of the event happened before.'

(https://www.facebook.com/groups/nalbeirakothabotora/permalink/4143312155719618)

In this example, the -k marked nominalized verb is followed by the postposition legi just like in early Assamese (see Example (17)). -k marked construction serves the function of purposive adverbial in the above example.

3.4. -k with a verb

Similar to the early Assamese, -k is still used to index third person subjects in some dialects (Kāmrupiyā and Barpetiā) of modern Assamese, as shown in the following example.

(36) ɔhuk kompandaɪ-or ɦoɹu apa bʰɛkla-i api
Ahuk compounder-gen younger boy Bhekla-erg girl

pɔl-e ni ita bɔɹpeta {pa-l-ak}.
hide-cnv bring.seq now Barpeta get-pfv-3
'The younger son of Ahuk compounder fled to Barpeta with a girl.'

(Sharma (2017)

It is lost entirely from the verbal indexes in the standard dialect of modern Assamese. The third person verbal index -ek (ēk) in early Assamese (see Example (18)) is preserved in two ways in the eastern (standard dialect) and western dialects (Kāmrupiyā, Barpetiā) of modern Assamese. It is used by dropping -k in the standard dialect, as in (37).

(37) ɔtul-e pani e-gilas {kʰa-l-e}
Atul-erg water one-glass eat-pfv-3
'Atul drank a glass of water.'

(Emille-CIIL Assamese Corpus: nmm133_CQPweb)

-ek (ēk) is used in the western dialects by changing the initial vowel e to a, as shown in (36) above. It is noted that the distribution of -ek in the western dialects is similar to early Assamese. It is confined to the context where it is always preceded by the perfective morpheme -l in both early and modern Assamese. On the contrary, -e in the standard dialect occurs next to the perfective morpheme -l (see (37))as well as next to the root morpheme[7] without being preceded by any tense/aspectual morpheme other than -l, as shown below.

(38) ɦi godʰuli ɦɔdai {pɔrh-e}
3inf.sg.m evening always read-3
'He always studies in the evening.'

(Self-elicited)

(39) ɦi godʰuli ɦɔdai {pɔrh-ib-ɔ}
3inf.sg.m evening always read-fut-3
'He will always study in the evening.'

(Self-elicited)

As shown in (39), the third person verbal index becomes -ɔ when it is preceded by the future morphmem -b (ib).

Apart from it, -k is also added to a verb to express a wish in modern Assamese, as presented in (40).

(40) kitap-kʰɔn iat-e {tʰak-ɔk}
book-clf.FLAT here-contrs stay-3.desire

[7] *-e* is realized as *-i* after the root morpheme that ends with a vowel other than a consonant.

'Let the book be here itself.'

(Self-elicited)

Conclusion

This study shows that starting from the genitive to the third person verbal index, -k is widely distributed with both nouns and verbs in early and modern Assamese. Even though by dropping some uses, such as the use of genitive with both ordinary nouns and nominalized verbs, the standard dialect of modern Assamese seems to show some progressive behavior, the western dialects such as Kāmrupiyā, Barpetiā seem to be still conservative retaining some of its older functions reflected in many early Assamese texts belonging to Pre-Sankardeva and Sankardeva period. It is yet to be studied whether the various functions discussed in this work can be considered as extended functions of a single morpheme or each function is carried by a separate morpheme, which is homophonous. It should be noted here that this study does not claim or consider -k as a single morpheme. It only surveys the various functions carried by a form, which shares an identical physical realization, on the basis of the synchronic data (both early and modern Assamese). Even though the diachronic aspect is considered to some extent for the sake of comparison of the functions between early Assamese and modern Assamese, it is less than enough to conclude anything like mentioned above. It certainly warrants more research which needs to incorporate cross-linguistic dimension along with the synchronic and diachronic distribution of the form to answer some relevant questions, such as whether -k is a single morpheme, which has multiple extended functions, or a homophonous morpheme; if it is a single morpheme, it is originated with which function.

References

Ananta Kandali. Kumar Haraṇ kābya. Assamese wikisource

Bez, Gitanjali. 2019. Possessive indexes in Assamese. Himalayan Linguistics 18(2). 1-34.

Rāy, Basantaranjan. (ed.). 1323 Bangabda. Caṇḍidāsar Sri Krishna kirtan. Calcutta: Visvakosha-Press

Hazarika, Parikshit. (ed.). 1992. Caryapada. Guwahati: Dalimi Prakashan (Assam, India).

Kakati, Banikanta. (1941). Assamese Its Formation and Development. Guwahati, LBS Publication, Assam, India.

Kandali, Madhav. Saptakānda Rāmāyana. H. N. Datta Barua (ed.), 1972.

Masica, Colin P. 1991. The Indo-Aryan Languages. Cambridge/Port Chester/Melbourne/Sydney, Cambridge University Press.

Saraswati, Hem. Prahlad Charit. https://as.wikisource.org/wiki.

Śastri, Haraprasād (ed.). Bangabda 1323. Hājār bacharer purāṇ bāngālā bhāṣāẏ bauddha gān o dohā.Calcutta: Bangiya Sahitya Parisad.

Sharma, Dhanjit, K. 2017. Bhelkar Bori. Abagunthan, vol. 5. Srimanta śangkardev. Kirtan Nāmghoṣā.

Sri Sri Mādhabdev. 1928 (published). Ādicarit. Guwahati: Guwahati New Press.

Abbreviations

1	First person	hon	Honorific
2	Second person	hum	Human
3	Third person	imp	Impersonal
acc	Accusative	impv	Imperative
add	Additive	inf	Inferior

advl	Adverbial	infv	Infinitive
all	Allative	loc	Locative
cft	Counterfactual	m	Masculine
clf	Classifier	neg	Negative
cnv	Converb	nmlz	Nominalizer
contrs	Contrastive	opt	Optative
dat	Dative	pf	Perfect
dem	Demonstrative	pfv	Perfective
dist	Distal	pl	Plural
erg	Ergative	poss	Possessive
excl	Exclamative	ques	Question
f	Feminine	seq	Sequential
fam	Familiar	sim	Simultaneous
fut	Future	tr	Transitive
gen	Genitive		

Jadavpur Journal of Languages and Linguistics
ISSN: 2581-494X

Personal pronouns in Kokborok and Dimasa

Samir Debbarma[1,1] & Kh. Dibyajyoti Barman[2]
[1]Tripura University, [2]Assam University

ARTICLE INFO

Article history:
Received 16 /07/2025
Accepted 06/10/2025

Keywords:
Tibeto-Burman,
Bodo-Garo,
Dimasa,
Kokborok,
personal pronouns

ABSTRACT

The present paper attempts to describe some of the aspects of morphosyntactic features of personal pronouns in Dimasa and Kokborok, the two Bodo-Garo languages spoken mainly in Assam and Tripura respectively. This study explores the personal pronominal system of Dimasa and Kokborok. The system does not differentiate between masculine and feminine in the personal pronouns i.e., or simply we can say that there is no gender distinction in personal pronouns of these two languages. Thus, the word bo refers to both he and she in both the languages. The first person form is suppletive in both the languages. While second and third person plurals are formed by suffixing to the singular forms.

Introduction

The Dimasa language (known as Grao Dima; ISO dis-639-3) and Kokborok known as Tripuri in Census of India (trp-639-3) are the two Bodo-Garo languages of Tibeto-Burman language family. According to the 2011 Census, the speakers of Tripuri (Kokborok) and Dimasa are 10, 11, 294 and 1,33,327 respectively. Kokborok is mainly spoken in Tripura, and the neighbouring states of Assam and Mizoram. It is also spoken in present day Bangladesh. Dimasa is mainly spoken in the states of Assam and Nagaland. Kokborok has eight distinct dialects namely Bru, Debbarma, Jamatia, Koloi, Murasing, Rupini, Tripura and Uchoi (Debbarma; 2001). With the recently added Noatia it becomes nine. The major dialects of Dimasa are Hasao, Hawar, Dembra, and Dijuha (mostly called Dijua).

The languages, Dimasa and Kokborok, come under the Barish (aka Bodo-Garo or Bodo-Koch) section of Baric (aka Sal or Bodo-Konyak-Jinghpaw) division of Tibeto-Burman language family (Shafer 1955: p.107). In Burling's (2003) classification, these two languages come under the Bodo-Koch branch of the Bodo-Konyak-Jinghpaw subgroup of the Tibeto-Burman sub-family of languages.

Figure 1: *Map of North East including Assam and Tripura*

[1] Corresponding Author Email ID: samirdebbarma85@gmail.com

Typological overview

Typologically, both Dimasa and Kokborok are agglutinating languages. Like many other South Asian languages, affixation, compounding, and reduplication are the three main word formation processes in these two languages. In affixal morphology, in comparison to prefixes, the occurrences of suffixes are more productive in both the languages. They have a rich set of classifiers like in other Bodo-Garo languages. In both the languages, the relative clause precedes the head noun whereas the adjectives, numerals and numeral classifiers usually follow the head noun. Verbs are marked for tense, aspect and mood. There is no subject-verb agreement in both the languages. Functionally, the nominal subject and nominal object precede the verb, and an adverb usually precedes the verb. More specifically, Dimasa and Kokborok are verb final languages, with dominant SOV word order.

Overview of Personal Pronouns in Bodo-Garo

The pronouns do not form a separate class different from the nouns, as far as affixal morphology is concerned. There is no gender distinction in the pronouns. Studies have shown some evidence that, in the modern BG languages, all the person distinctions made are not consistently found across the family, nor are all the forms cognate (Wood: 2008). Table 1 demonstrates the closeness in form for the 1s, 2s and 1pl pronouns, which easily reconstruct to the original PBG pronominal forms. They also reflect the Proto-Tibeto-Burman system (Matisoff: 2003; Jacquesson: 2005). Following the sound correspondences proposed by Joseph & Burling (2006), it can be reconstructed as *aŋ, *naŋ and *ciŋ, respectively.

	Boro	**Deuri**	**Dimasa**	**Garo**	**Atong**	**Ruga**	**Kokborok**	**Rabha**	**Tiwa**	**PBG**	**PTB**
1S	aŋ	ã	aŋ	aŋ(a)	aŋ~aŋa	aŋ	aŋ	aŋ	aŋ	* aŋ	* aŋ
2S	nɯŋ	no	niŋ	na?(a)	naŋ? ~na?a	naŋ	nuŋ	naŋ	na	*naŋ	*naŋ
3S	bi	ba	bo	bo	ge?theŋ~ de?theŋ	ua	bo	u a	pe		
1P	jɯŋ	jou	jiŋ	na?ching an?ching		niŋ	cuŋ	ciŋ	chiŋ	*ciŋ	*ciŋ
2P	nənsə r	lou	nisi	na?soŋ	naŋ?t-m	nanu ŋ	norok	naroŋ	nabur		
3P	bisər	bau	bonsi	bi-soŋ	ge?theŋtheŋ it-m ut-m~ -t-m	inuŋ	bo-rok	i-roŋ uroŋ	pibur		

Table 1: *Personal Pronoun across Bodo-Garo*

The third person pronouns slightly diverge within the family. There is a strong correspondence across the family with respect to the /bi/- initial forms, which suggests pre-PBG origin. This PBG pronoun reconstructs as *bi (Joseph & Burling 2006).

Personal pronouns in Dimasa and Kokborok

Pronouns in Dimasa and Kokborok are mostly noun-like, and even considered to be a sub-class of nouns, however, they do differ from other nouns in a few respects; they rarely take any sort of modifiers, and they can replace or substitute nouns. There is no inclusive/exclusive distinction (Payne 1997) of the personal pronouns in these two languages. Personal pronouns in Dimasa and Kokborok differentiate three persons and two numbers. In personal pronouns, there is no gender distinction in both the languages. They can be suffixed with various morphemes like plural

markers, case markers, particles etc. The Singular personal pronouns, in Dimasa and Kokborok, are free forms. The second and third person singular forms can be pluralized by suffixing the morpheme -rok in Kokborok, and -si in Dimasa. However, like other Bodo-Garo languages, the first person plural pronouns are suppletive forms in Dimasa and Kokborok.

First Person

In Kokborok and Dimasa, first person singular is a closed monosyllabic word consisting of VC syllabic pattern. In both the languages first person singular personal pronoun begins with central low unrounded vowel a and ends with velar nasal consonant ŋ, that is aŋ. The nasal final in first person pronoun is one the typical features of Bodo-Garo languages as shared by Kokborok and Dimasa and it corresponds to both the Proto-Bodo-Garo (PBG) and Proto-Tibeto-Burman (PTB) 1SG personal pronoun *aŋ (Matisoff, 2003, Jacquesson, 2005).

Like other modern Bodo-Garo languages, the first person pronoun in Kokborok and Dimasa the languages manifest a special plural form. In both the languages the first person plural personal pronouns are suppletive forms. The suppletive forms are monosyllabic closed forms, and both the forms share the same syllabic pattern, i.e., CVC pattern. At the onset position, the first person plural pronoun manifests homorganic sound, i.e., alveo-palatal sound. However, it contrasts in voiced and voiceless, i.e., c in Kokborok and z in Dimasa. Both the languages have velar nasal consonant ŋ in coda position. In the nucleus position, the first person plural personal pronoun manifests high vowels: u and i, with back and front contrast, in Kokborok and Dimasa, respectively.

Person	Personal pronouns		
1st	Language	Singular	Plural
	Kokborok	aŋ	ziŋ
	Dimasa	aŋ	cuŋ

Table 2: *First person of Dimasa and Kokborok*

Second Person

In both languages, the second person singular is a monosyllabic form, consisting of a CVC syllabic pattern. Second person singular is closed syllable in Kokborok and Dimasa, the velar nasal consonant ŋ occurs in coda position. In both the languages, second person singular begins with the alveolar nasal n, such as nuŋ (Kokborok) and niŋ (Dimasa). Here in the case of second person singular pronoun too, it corresponds to both the Proto-Bodo-Garo (PBG) and Proto-Tibeto-Burman (PTB) 2SG personal pronoun *naŋ (Benedict 1972, Matisoff 2003, Jacquesson, 2005).

In Kokborok second person plural is disyllabic word, plurality is made by means of affixing the plural morpheme -rok to second person singular pronoun no (derived from nuŋ). In Dimasa, the second person plural is a disyllabic word, formed by affixing the plural morpheme -si (only in case of personal pronouns) to second person singular pronoun ni derived from niŋ. The second person plural is closed syllable in Kokborok whereas in Dimasa it is open. Though it begins with alveolar nasal in both the languages, but Kokborok has consonant (voiceless velar plosive k) in its final whereas in Dimasa it is high vowel (unrounded close high i).

The second person pronouns in Kokborok and Dimasa are summarized in the following table 3.

Person	Personal pronouns		
2nd	Language	Singular	Plural
	Kokborok	nuŋ	no-rok
	Dimasa	niŋ/nuŋ	ni-si

Table 3: *Second person of Dimasa and Kokborok*

Third Person

In Kokborok and Dimasa the third person singular pronoun is a monosyllabic word consisting of CV syllabic pattern. Begins with voiced bilabial stop b- followed by the mid back rounded vowel -o. As in many other Tibeto-Burman languages, Kokborok and Dimasa also don't make any gender distinction in pronouns particularly in third person pronouns as we find in Indo-European languages like Assamese, Sylheti Bengali, English, etc.

In Kokborok, the 3rd person plural pronoun is made by means of suffixing the same plural morpheme -rok, which is found in second person plural and other nouns other than kinship terms. Similarly in Dimasa too, the 3rd person plural is made by suffixing the same plural morpheme -si which is found in the 2nd person singular pronoun.

It is worth mentioning here that the plural morpheme -si cannot directly attach to the third person singular pronoun in Dimasa. Morphophonemically, the high front vowel i of genitive marker -ni is being deleted when it is followed by the same vowel. As a result, the underlying form of bonisi became bonsi. Third Personal Pronouns in Dimasa are illustrated below in table 4:

Language	3rd Person Personal Pronoun	
	Singular	Plural
Kokborok	bo	bo-rok
Dimasa	bo	bun-si

Table 4: *Third person in Dimasa and Kokborok*

Pronominal prefixes

Dimasa and Kokborok pronouns do have phonologically reduced forms in the form of pronominal prefixes that can be prefixed to kinship nouns and possessive forms.

Dimasa

(1) a-da ‘elder brother’
1-elder brother

(2) a-ni ‘my/mine’
1-gen

(3) nə-da ‘your elder brother’ (senior to junior)
2-elder brother

(4) bi-gim ‘his brother-in-law’
3-brother-in-law

Kokborok

(5) a-ta ‘elder brother’
1-elder brother

(6) a-ni ‘my/mine’
1-gen

(7) a-ta ‘elder brother’
1-elder brother

(8) a-ni ‘my/mine’
1-gen

Case marking on pronouns

Case markers can be suffixed to pronouns just as they can be suffixed to nouns. The declension of personal pronouns is illustrated in Table 5.

Language	Person	Accusative	Dative	Genitive	Comitative	Locative
Dimasa	1sg	aŋ-k^{h}e ‘me’	an-t^{h}ene ‘to me’	a-ni ‘my’	aŋ-zaŋ ‘with me’	a-ni-ha ‘in me’
	2sg	niŋ-k^{h}e ‘you’	nin-t^{h}ene ‘to you’	ni-ni ‘your’	niŋ-zaŋ ‘with you’	ni-ni-ha ‘in you’
	3sg	bo-k^{h}e ‘him/her’	bun-t^{h}ene ‘to him/her’	bu-ni ‘him/her’	bo-zaŋ ‘with him’	bo-ni-ha ‘in him/her’
	1pl	ziŋ-k^{h}e ‘me’	zin-t^{h}ene ‘to me’	zi-ni ‘my’	ziŋ-zaŋ ‘with us’	zi-ni-ha ‘in me’
	2pl	nisi-k^{h}e ‘you’	nisi-ne ‘to you’	nisi-ni ‘your’	nisi-zaŋ ‘with you’ (pl)	nisi-ni-ha ‘in you’
	3pl	bunsi-k^{h}e ‘to him/her’	bunsi-ne ‘to him/her’	bunsi-ni ‘his/her’	bunsi-zaŋ ‘with them’	bunsi-ni-ha ‘in him/her’
Kokborok	1sg	a-no ‘me’	a-no ‘to me	a-ni ‘my’	aŋ-bai ‘with me’	
	2sg	no-no ‘you’	no-no ‘to you’	ni-ni ‘your’	nuŋ-bai ‘with you’	
	3sg	bo-no ‘his/her’	bo-no ‘to him/her’	bo-ni ‘him/her’	bo-bai ‘with him’	
	1pl	cu-no ‘me’	cu-no ‘to me	ci-ni ‘my’	cuŋ-bai ‘with us’	
	2pl	norok-no ‘you’	norok-no ‘to you’	norok-ni ‘your’	norok-bai ‘with you’ (pl)	
	3pl	borok-no ‘to him/her’	borok-no ‘to him/her’	borok-ni ‘his/her’	borok-bai ‘with them’	

Table 5: *case marking of personal pronouns of Dimasa and Kokborok*

Morphophonemics in Personal Pronouns

Morphophonemic changes take place in the personal pronouns of Dimasa and Kokborok when combined with case marking, plural markers and in pronominal prefixes as discussed below.

Deletion

Dimasa

In Dimasa the velar nasal /ŋ/ is being dropped when it is followed by alveolar nasal sound /n/ as shown in table 6.

Underlying form	Reduced form	Gloss
aŋ-ni	ani (1s-gen)	my
ziŋ-ni	zini (1pl-gen)	our
niŋ-ni	nini (2s-gen)	you
niŋ-si	nisi (2- pl)	you (pl)

Table 6: *deletion in Dimasa*

Kokborok

Similar morphophonemic changes also occur in Kokborok where the alveolar nasal is being dropped when it is followed by alveolar sound as illustrated in table 7.

Underlying form	Reduced form	Gloss
aŋ-ni	ani (1s-gen)	my
cuŋ-ni	cini (1pl-gen)	our
nuŋ-ni	nini (2s-gen)	your
nuŋ-rok	norok (2-pl)	you-pl
aŋ-no	a-no (1-acc/dat)	to me
nuŋ-no	no-no (2-acc/dat)	to you

Table 7: *deletion in Kokborok*

Thus, deletion of velar nasal sound, especially in first and second person personal pronouns, is one of the corresponding features, found in both the languages.

Alveolarization

In Dimasa, the velar nasal ŋ became alveolar n when it is followed by the alveolar sound t as shown in (9) and (10). It is worth mentioning that this type of alternation is found when case markers are attached to first and second person singular personal pronouns. This process is not found in Kokborok.

(9) aŋ-t^{h}ene > an-t^{h}ene (1-dat)

(10) niŋ-t^{h}ene > nin-t^{h}ene (2-dat)

Vowel harmony in plural formation and Case marking

Vowel harmony is also seen in plural formation. The pronominal prefix harmonizes with the root of the plural suffix in Dimasa (11) and Kokborok (12). The pronominal prefix in Kokborok also harmonizes with the accusative and dative suffix -no as in (13).

(11) niŋ-si > nisi (2-pl)

(12) nuŋ-rok > norok (2-pl)

(13) nuŋ-no > no-no (2sg-dat/acc)

Vowel harmony in Possessive construction

In Kokborok, the possessive or genitive marker triggers vowel harmony of the first person plural, second and third person singular pronouns. In other words, the vowels of pronominal base forms harmonize with the vowel of the genitive morpheme as illustrated in (14) – (16).

(14) ci-ni nok kotor
3pl-gen house big
'Our house is big.'

(15) ni-ni nok nait^{h}ok
2sg-gen house beautiful
'Your house is beautiful'

(16) bi-ni mokol kaham
3sg-gen eye good
'His eye is good.'

In Dimasa, especially in its Hawar[2] variety, regressive vowel harmony is seen in the first person plural and second person singular personal pronouns. The vowels of pronominal base forms harmonize with the vowel of the genitive morpheme as illustrated in (17) and (18).

(17) zuŋ-ni > zi-ni nolai de-bi
3pl-gen 1pl-gen village big
'Our village is big.'

(18) nuŋ-ni > ni-ni no bra-ha
3pl-gen 2sg-gen house where-loc
'Where is your house?'

Vowel harmony in Pronominal prefixes

Kokborok

The second person singular pronoun nuŋ 'you' has four forms of pronominal prefixes na-, ni-, no-, and nu-. In other words, the vowels of pronominal prefixes nV- harmonize with the vowels of nominal roots particularly kin nouns as illustrated in (19) – (24).

(19) na-hamɟuk
2sg-daughter-in-law 'your daughter in-law'

(20) na-hanok
2sg-sister 'your younger sister'

(21) ni-kiciŋ
2sg-friend 'your friend'

(22) no-toi
2sg-small maternal aunty 'your small maternal aunty'

(23) nu-cu
2sg-grandfather 'your grandfather'

(24) nu-kumui
2sg-brother-in-law 'your brother-in-law'

Similarly, the third person bo 'he' has four forms of pronominal prefixes ba-, bi-, bo-, and bu- as illustrated in the following examples (25) – (29).

(25) ba-hamɟuk
3sg-daughter-in-law 'his daughter in-law'

(26) ba-hanok
3sg-sister 'his younger sister'

(27) bi-kiciŋ
3sg-friend 'his friend'

(28) bu-cu
3sg-grandfather 'his grandfather'

(29) bu-kumui
3sg-brother-in-law 'your brother-in-law'

Dimasa

As like in the Kokborok, the second person singular pronoun niŋ 'you' has four forms of pronominal prefixes na-, ni-, ne-, and nu- on the basis of the nucleus of the nominal roots as in (30) – (37).

(30) na-(h)andao

[2] Found mostly in southern most districts of Assam: Cachar, Karimganj and Hailakandi.

2sg-youger sister ‘your younger sister’

(31) ni-bi
2sg-elder sister ‘your elder sister’

(32) ni-gim
2sg-brother-in-law ‘your brother in-law’

(33) ne-de
2sg-maternal uncle ‘your maternal uncle’

(34) na-mao
2sg-maternal aunty ‘your maternal aunty’

(35) nu-zu
2sg-grandfather ‘your grandfather’

(36) na-dai
2sg-grandmother ‘your grandmother’

(37) ni-di
2sg-paternal aunty ‘your paternal uncle’

Similarly, the third person bo ‘he’ has four forms of pronominal prefixes ba-, bi-, be-, and bu- as illustrated in the following examples (38) – (44).

(38) ba-handao
3sg-younger sister ‘his younger sister’

(39) ba-dai
3sg -grandmother ‘his grandmother’

(40) ba-mao
3sg-maternal aunty ‘his maternal aunty’

(41) bi-gim
3sg- brother-in-law ‘his brother-in-law’

(42) bu-zu
3sg- grandfather ‘his grandfather’

(43) bi-di
3sg-paternal uncle ‘his paternal uncle’

(44) be-de
3sg-his maternal uncle ‘his maternal uncle/paternal aunty’

Syntax of personal pronouns

Personal pronouns function as subjects and objects (Kroeger 2005) and they modify nouns as genitive pronouns. The pronouns function as subject arguments or subject complement. It is observed that the nominative form of pronouns is unmarked in Kokborok as in (45) and (46) in Dimasa too it is unmarked as in (47) and (48).

(45) Kokborok:
aŋ-ø bo-bai hat^{h}i-o malai-k^{h}a
I-nom he-commarket-loc meet-pst
‘I met with him in the market.’

(46) Kokborok:
bo-rok-ø cuŋ-bai-se p^{h}ai-k^{h}a
he-pl-nom we-com-emp come-pst
‘They came with us.’

(47) Dimasa:
aŋ-ø bo-zaŋ hat^{h}ai-ha goɾon-ba
I-nom he-commarket-loc meet-pst
'I met with him in the market.'

(48) Dimasa:
bun-si-ø ziŋ-zaŋ pai-pa-ba
he-pl-nom we-com come-attach-pst
'They came with us.'

Like many other TB languages, Kokborok and Dimasa don't have distinct oblique forms of personal pronouns. However, the oblique forms of personal pronouns are formed by postponing case markers to the corresponding personal pronoun as illustrated in the following examples.

Kokborok:
(49) aŋ bo-no nuk-k^{h}a
I he-acc see-pst
'I saw him.'

Dimasa:
(50) bo-k^{h}e p^{h}ai-ma-ne t^{h}i-ha
3-acc come-inft-dat tell-pol
'Tell her to come.'

Kokborok:
(51) tini nuŋ aŋ-ni bagui amotui tubu-di
today you I-gen for pineapple bring-imp
'Today you bring pineapple for me.'

Dimasa:
(52) dakna niŋ an-t^{h}ane laimoɾi labu-hi ri
tomorrow you I-dat pineapple bring-conj give
'Tomorrow you bring pineapple for me.'

Prodrop

Like other Bodo-Garo language, Dimasa and Kokborok languages also manifest pro-drop feature, i.e., while the context is clear to the addressee the pronouns of the clause or sentence, that is the subject argument is being dropped without any distortion of the sentence semantically (53) and (54). It is also observed that if the subject arguments are nouns then they are obligatorily used in the sentence, that is, the subject argument cannot drop in the sentence as illustrated in (55) and (56).

Dimasa:
(53) p^{h}ɾoŋ-ha p^{h}ai-naŋ
morning-loc come-fut
'(I) will come in the morning.'

Dimasa:
(54) biman p^{h}ɾoŋ-ha p^{h}ai-naŋ

Biman morning-loc come-fut
'Biman will come in the morning'
Kokborok:
(55) $p^huŋ$-o p^hai-nai
Morning-loc come-fut
'(I) will come in the morning.'
Kokborok:
(56) biman $p^huŋ$-o p^hai-nai
Biman morning-loc come-fut
'Biman will come in the morning'

Conclusion

It can be concluded that personal pronouns in Dimasa and Kokborok are free in both singular and plural forms. They have no inclusive/exclusive distinction and they do not mark gender. Some pronouns have phonological conditioning when suffixed with case markers. Just like nouns, they can function as subjects, objects and oblique objects and also take nominative (unmarked), accusative and dative case markers respectively. The absence of nominative markers in both Kokborok and Dimasa is a typical feature in contrast to Garo and Boro.

Both in Kokborok and Dimasa the first person plural pronoun is suppletive in form which is not common in other Tibeto-Burman languages. The pluralization of second and third person pronouns in both the languages are done by suffixes. The generic plural marker -rok is used in Kokborok to pluralize the second and third person respectively. While in Dimasa it is marked by -si (only in personal pronouns).

Morphophonemic changes occur in both the languages. The velar nasal is being dropped when it is followed by alveolar sound, in both the languages. In Dimasa, the velar nasal becomes alveolar nasal when it is followed by the alveolar consonant. The vowels of personal pronouns, in Kokborok and Dimasa, harmonize with vowels of genitive morpheme and plural morpheme. Vowel harmony doesn't apply to the first person pronoun as it begins with vowel.

References

Barman, Kh. D. (2018). Pronouns in Dimasa. Unpublished M.Phil. Dissertation, Assam University, Silchar.

_____ (2023). Marked Pronouns in Three Barish Languages. In Language in India, Vol. 23:10

Benedict, P.K. (1972). Sino-Tibetan A Conspectus. USA: Cambridge University.

Bhatt, D.N.S. (2004). Pronouns. Oxford: Oxford University Press.

Burling, Robbins. (2003). "The Tibeto-Burman Languages of North-Eastern India". In Graham Thurgood and Randy J. LaPolla Eds., The Sino-Tibetan Languages. London and New York: Routledge. 169-191.

Census of India. (2011). Series 1: Language India and States. New Delhi: Registrar General of India and Census Commissioner.

Debbarma, Binoy. (2001). Concise Kokborok-English-Bengali dictionary. Khumulang: Language Wing.

Dixon, R.M.W. (2010). Basic Linguistic Theory: Grammatical Topic. Volume 2. Oxford/New York: Oxford University Press.

Jacquesson, F. (2005). Le Deuri: Langue Tibéto-Birmane d'Assam. Peeters: Leuven.

Jaquesson, F. (2007). Kokborok Grammar: Agartala Dialect. Agartala: KOHM.
Jacquesson, F. (2008). A Dimasa Grammar. (PDF) A Short Dimasa Grammar | Francois Jacquesson - Academia.edu
Joseph, U.V. and Robbins Burling. (2006). The Comparative Phonology of the Boro Garo Languages. Mysore: Central Institute of Indian Languages.
Karapurkar, Pushpa P. (1979). Kokborok Grammar. Mysore: Central Institute of Indian Languages.
Matisoff, James A. (2003). Handbook of Proto-Tibeto-Burman: system and philosophy of Sino-Tibetan reconstruction. Berkeley & Los Angeles: University of California Press.
Payne, Thomas E. (1997). Describing Morphosyntax: A Guide for Field Linguists. Cambridge: Cambridge University Press.
Shafer, Robert. (1955). "Classification of the Sino-Tibetan languages". Word. 11: 94-111.
Wood, D.C. (2008). An Initial reconstruction of Proto-Boro-Garo. Unpublished M.A. Dissertation, University Of Oregon, USA.

Abbreviation

1	first person	INFT	Infinitive
2	second person	LOC	Locative
3	third person	NOM	Nominative
ACC	Accusative	PBG	Proto-Bodo-Garo
BG	Bodo-Baro	PL	Plural
C	Consonant	POL	Polite
COM	Comitative	PST	Past
CONJ	Conjunctive	PTB	Proto-Tibeto-Burman
DAT	Dative	SG	Singular
EMP	Emphatic	SOV	Subject-Object-Verb
FUT	Future	TB	Tibeto-Burman
GEN	Genitive	V	Verb

Jadavpur Journal of Languages and Linguistics
ISSN: 2581-494X

Individuative and Generic Referential NPs in Tutsa

Monjushna Sudhakaran[1] *& Gautam K. Borah*
Tezpur University, Assam

ARTICLE INFO

Article history:
Received 27/08/2025
Accepted 10/09/2025

Keywords:
Tutsa,
Individuative reference,
Indefinite quantifiers,
definite collective markers,
generic reference

ABSTRACT

This paper describes the grammar and semantics of individuative and generic referential noun phrases in Tutsa, a Tibeto-Burman language spoken primarily in Arunachal Pradesh, India. Our analysis reveals that Tutsa employs a system where bare nouns can function as NPs with contextually determined indefiniteness, definiteness, and also number. Similarly, a numerically quantified NP in Tutsa is either contextually indefinite or definite; a non-numerically quantified indefinite NP is marked by a non-numerical quantifier that is inherently indefinite, while its definite counterpart is marked by a non-numerical quantifier that is inherently definite, which we designate as a definite collective marker. For reference in physical space, Tutsa uses two demonstratives, which are simultaneously singular and plural. For generic reference, Tutsa utilizes the bare noun to represent the entire class, or a noun marked by the definite collective marker -lɔŋ, resulting in a definite plural NP that metonymically represents the entire class. This initial description contributes to the documentation of understudied Tibeto-Burman languages and provides a groundwork for further comparative research.

Introduction

Effective communication relies on how speakers establish a shared understanding of the entities being discussed. This is achieved through the linguistic mechanism of reference, where noun phrases (NPs) are used to identify and talk about specific instances of a kind of thing, as well as the kind itself. For instance, a speaker might use one form to point out 'this boy' or refer to 'that boy', while another form might be used when talking about 'boys' in general. Reference is thus a human act that imbues linguistic expressions with meaning. To quote Hanks (2019: 11): "People refer to things, and their doing so makes it possible for expressions to refer. Linguistic expressions, after all, are just collections of sounds or marks, with no intrinsic ability to stand for anything else. It is because of the way people use these sounds and marks that they come to refer to things outside themselves." This paper examines the grammatical structures and meanings associated with these different ways of using noun phrases in Tutsa, aiming to describe the system of what is commonly understood as individuative reference (i.e. referring to specific instances of a kind) and generic reference (i.e. referring to the entire kind or class) within the language.

Tutsa is a tonal Tibeto-Burman language with two tones (high and low), and it is classified within the larger Sino-Tibetan language family. More specifically, it is part of the Tangsa group, which is a cluster of closely related languages. Tutsa, which is genetically linked

[1] Corresponding Author Email ID: monjushna@gmail.com

to major languages such as Tibetan, Burmese, and other Sino-Tibetan languages spoken across Asia, is primarily spoken in the northern and southern districts of Changlang and eastern Tirap in Arunachal Pradesh, India, with some speakers also residing in parts of the Tinsukia district of Assam.

With approximately 25,000 speakers (this figure is based on the 2001 census data), about half of whom are monolingual, Tutsa faces significant challenges to its vitality. The language has no established written literature; its rich oral traditions, including songs, folk tales, and stories, exist primarily in colloquial form. Speakers currently use the Roman script. The influence of dominant regional languages like English, Hindi, and Assamese contributes to a lack of cohesion in language use among its speakers, with the younger generation increasingly drawn to these more widely used languages.

According to the UNESCO Report, the Tutsa language is classified as EGIDS Level 6B (Threatened), which indicates that while it is used for face-to-face communication across all generations, it is experiencing a decline in its number of users. The Tutsa were previously recognized as part of the Nocte tribe but were later classified as a separate tribe around 2000 (Morey 2017). Their language, which has an ISO code (ISO 639-3: **tvt**), is also listed in Ethnologue (Lewis et al. 2014) as 'Naga, Tutsa' (also see Hammarström, Harald et al 2025; Matisoff, James A. 1996).

No significant comprehensive linguistic studies of Tutsa's grammar or structure are available to date. While Rehung (1992) produced an early, 124-page account of Tutsa grammar, its approach does not align with modern linguistic analysis, and it offers no significant explanation of reference. Similarly, Rao (2010) offers a two-part work on the Tutsa tribe, focusing on their origin, migration, and history, but does not provide a grammatical study of the language itself. Additionally, a 200-page publication by the Vivekananda Kendre Institute of Culture in Guwahati outlines the traditional life of the Tangsa and Tutsa tribes. While it emphasizes their animistic beliefs, agricultural practices, and unique social customs, it does not include any grammatical analysis of the language.

A research paper by Konch (2021) offers a study on the nominal inflection of the Tutsa language, and Bishakha Das's (2021) article is a comparative work on the auxiliary verbs in Nocte, Khappa, Ollo, and Tutsa. The data used for this analysis has been collected from native Tutsa speakers through our fieldwork in their villages of Arunachal Pradesh and Assam. This collection forms a vital part of our ongoing project to compile a comprehensive descriptive grammar of this critically understudied and endangered language. This paper's description of reference in Tutsa therefore offers a crucial initial contribution to the documentation of this valuable linguistic heritage.

The remainder of this paper is structured as follows: Section 2 begins by outlining the theoretical framework of reference and then introduces the two main types investigated: individuative and generic. Section 2.1 provides a detailed analysis of the ways of individuative reference in Tutsa, showing how the bare noun serves as an individuative referential expression; the numerically and non-numerically quantified NPs as referential expressions; the demonstrative-marked NPs as referential expressions. Section 2.2 focuses on the strategies for achieving generic reference in Tutsa through the use of the bare noun and non-numerically quantified NP as a metonymic expression for the entire class. Finally, Section 3 summarizes our key findings and identifies promising avenues for future research into Tutsa grammar and its broader typological connections within Tibeto-Burman languages.

The remainder of this paper is structured as follows. Section 2 begins by outlining the theoretical framework of reference before introducing the two main types investigated: individuative and generic. Section 2.1 provides a detailed analysis of individuative reference in Tutsa, showing how the bare noun, numerically and non-numerically quantified noun phrases (NPs), and demonstrative-marked NPs all function as individuative expressions. Section 2.2 focuses on the strategies for achieving generic reference in Tutsa through the use of the bare noun and non-numerically quantified NPs as metonymic expressions for an entire class. Finally, Section 3 summarizes our key findings and identifies promising avenues for future research into Tutsa grammar and its broader typological connections within Tibeto-Burman languages.

Types of Reference in Tutsa

Reference, a core linguistic mechanism, enables speakers to guide hearers toward a mutual understanding of the entities being discussed. Effective communication relies on a speaker's capacity to establish a shared focus, ensuring the listener identifies the precise item intended. As an illustration, the Tutsa example ŋa-é lɔrí esít rehtàŋ ('ŋa-é 'I-erg; lɔrí 'car'; esít 'one'; rehtàŋ ' 'bought') 'I bought one car' demonstrates this; here, lɔrí esít 'one car' specifically refers to a single vehicle. Such an act of reference is crucial for grounding particular instances within the ongoing discourse.

Our observations of Tutsa reveal two main types of reference: Individuative and Generic. Individuative reference involves referring to a single instance of a class or kind, illustrated by the NP erá nɔ̀k 'that house' in (1), which points to a specific dwelling. Generic reference, conversely, involves the speaker referring to an entire kind of thing, as shown in (2) with the bare noun d͡ʒuŋwɔ meaning 'crane' referring to the whole category of these birds.

(1) erá nɔ̀k áh d͡ʒàsemɔ̀ŋ
that house be beautiful
'That house is beautiful.'

(2) d͡ʒuŋwɔ áh epùŋ
crane be white
'Cranes are white.'

Individuative reference

In a language like English, individuative reference is typically marked, which is why a sentence like *'He bought car yesterday' is grammatically incorrect. The noun 'car' needs a determiner like 'a' to make it an acceptable indefinite NP, or 'the' to make it a definite NP. This requirement highlights how crucial these dedicated words are for conveying individuative meaning in English. Without them, a speaker cannot specify a single entity.

However, languages like Tutsa demonstrate that individuative reference is not always marked by a dedicated word. Instead, the context of the conversation and the shared knowledge between speakers are what allows a bare noun to be understood as referring to a specific individual. This in turn shows that the mental act of isolating a single entity remains the same, but the linguistic tools used to achieve that act can vary greatly from one language to another. Thus, depending on the context involved, the same NP in Tutsa may serve as an indefinite or definite NP, as we will see below.

Bare nouns as referential expressions

In Tutsa, we find that the bare noun form can also function as an NP, exhibiting a range of possible interpretations as a referential expression. Thus, depending on the context involved, it may have a generic reference as in (2) above; an indefinite singular/plural; and a definite singular/plural, as in, (3) below. This characteristic, where nouns in their basic form can act as NPs with contextually flexible reference, is not unique to Tutsa. As shown by Borah (2008; 2011), Assamese is another language where bare nouns similarly function as NPs, further illustrating this typological pattern.

(3) wén wàŋ-hala
guest come-prog
'A guest is coming/The guest is coming.'
'Some guests are coming/The guests are coming.'

(4) d͡ʒuŋwɔ́ áh d͡ʒuŋ naŋ tʰa:m-ta
crane be lake loc land-pst
'A crane/The crane has landed at the lake.'
'Cranes/The cranes have landed at the lake.'

In Tutsa, the context of a statement determines whether a bare noun is interpreted as generic or individuative. In Example (2), the bare noun d͡ʒuŋwɔ́ 'crane' is used in a statement that describes a general, permanent characteristic of the entire species (i.e. that cranes are white). This is a statement of universal fact, not a comment about a specific bird. The context of a general truth prompts the listener to understand the noun as representing the whole class of cranes, not a particular one.

By contrast, Example (3) uses the bare noun wén 'guest' in a statement about a specific, real-world event: someone is arriving. This cannot be a general statement about all guests in the world, as that would be illogical. The context of a unique, specific event, i.e. a guest's arrival, forces an individuative interpretation. The listener understands that the noun refers to one or more particular guests who are currently present or on their way. The specific number and identity are determined by this immediate, real-world context rather than by the language's grammar itself.

Similarly, Example (4) is a statement about a specific, real-world event: a crane landing at a lake. Because this is not a general truth about all cranes or all lakes, the context of the sentence forces an individuative interpretation. The bare noun d͡ʒuŋwɔ́ 'crane' cannot be understood as referring to the entire class of cranes, i.e. 'All cranes have landed at the lake', because such a statement would be illogical. The exact meaning, i.e. whether it is 'a crane,' 'the crane,' 'some cranes,' or 'the cranes', is determined by the immediate context. For example, if a speaker and listener see a single, unknown crane land, they would understand it as 'a crane.' If they were talking about a specific, known crane, they would understand it as 'the crane.' This example perfectly illustrates how in a language like Tutsa, the bare noun's meaning is highly flexible and relies entirely on the pragmatic context of the conversation.

Numerically quantified NPs as referential expressions

The numerically quantified NP seems to be a linguistic universal because quantifying objects is a fundamental human need. All cultures and languages require a way to refer to specific amounts of things, whether it is for trading, building, or simply counting.

From a cognitive perspective, the ability to count and distinguish between singular and plural items is a basic part of human thought. Therefore, languages have evolved grammatical and lexical mechanisms, numbers and quantifiers, to express these essential concepts. While the way different languages mark this may vary, the core function of a numerically quantified NP is present everywhere because it is tied to a universal human need to interact with the world in terms of number.

Thus, Tutsa employs numerically quantified NPs, as in (5) and (6). A numerically quantified NP is either indefinite or definite. However, our analysis of Tutsa reveals that the numerically quantified NP does not employ specific grammatical markers to distinguish between indefiniteness and definiteness. Examples (5) and (6) provide evidence for this:

(5) ŋa-é minusá wazɔ́m kaptup-tà-ŋ. minusá wazɔ́m
I-erg girl three see-pst-1 girl three
áh ʤàsemɔ̀ŋ.
be beautiful
'I saw three girls. The three girls were beautiful.'

(6) ŋa tàŋnáŋ mayén esít eʤé. mayén áh
I have cow one exists. cow be
élɔ̀ŋ
big
'I have one cow. The cow is big.'

The first sentence of example (5) introduces minusá wazɔ́m 'three girls'. This is an indefinite reference, meaning the listener does not know which specific girls the speaker is talking about yet.

In the second sentence of (5), the same NP minusá wazɔ́m refers back to the same three girls that were just introduced. This time, the reference is definite because the speaker and listener now share knowledge about which specific girls are being discussed.

Thus, Tutsa does not add a word like 'the' to make it definite: the phrase minusá wazɔ́m 'three girls' is exactly the same in both sentences.

Similarly, the first sentence of (6) introduces mayén esít 'one cow.' In the second sentence, mayén 'cow', refers back to the same cow. The reference is now definite because the listener knows which specific cow is being talked about. The definite meaning comes from the shared context established by the first sentence.

Note that the word esít 'one' is dropped in the second sentence of Example (6) because it is no longer needed. The purpose of esít in the first sentence is to introduce a new, indefinite singular noun 'one cow'. Once that cow has been introduced and is part of the shared knowledge between the speaker and the listener, it becomes a definite referent. One can then refer to it as just mayén 'the cow', because its singular number is already understood from the previous context. The explicit mention of 'one' is no longer necessary to convey the meaning.

It is clear from (5) and (6) that numerically quantified NPs in Tutsa, as said, do not employ any grammatical marker to distinguish between indefiniteness and definiteness. In the initial sentences of (5) and (6), the NPs minusá wazɔ́m 'three girls' and mayén esít 'one cow', respectively, appear without any overt indication of being indefinite. However, their subsequent

occurrence in juxtaposed sentences with definite reference suggests that definiteness in these constructions is contextually derived.

Non-numerically quantified NPs as referential expressions

Languages have also non-numerically quantified NPs because people often refer to quantities imprecisely, using general terms instead of exact numbers. We do not always need to count things exactly in daily life (see Borah 2008).

In a language like English, speakers use words like 'some,' 'many,' 'few,' and 'all' to refer to groups of things without knowing or stating the precise number. These non-numerical quantifiers are essential for everyday communication, where general estimations are often sufficient and more efficient than precise counting. The existence of these types of NPs across languages shows that the human need for imprecise quantification is just as fundamental as the need for precise counting.

Thus, Tutsa employs non-numerically quantified NPs, as in (7) and (8) below. Our analysis of Tutsa reveals that the use of non-numerical quantifiers like héntékték 'many' and ekʰɔkʰɔ 'a few' in modifying a noun (as in the first sentences, respectively, of 7 and 8) consistently results in noun phrases that are interpreted as indefinite. This suggests that in Tutsa, these quantifiers contribute to the indefiniteness of the NP, i.e. they are inherently indefinite[2].

(7)

ŋa-é	*d͡ʒuŋwɔ́*	*héntékték*	*kaptup-tà-ŋ.*	*d͡ʒuŋwɔ́-lɔŋ*	*áh*
I-ERG	crane	many	see-PST-1.	crane-DCM	be
élɔ́ŋlɔ́ŋ	*aŋ-tà*				
big	be-PST				

'I saw many cranes. The cranes were big.'

(8)

ŋa tàŋnáŋ	*mayén*	*ekʰɔkʰɔ*	*edʒé.*	*mayén-*
I have	cow	a few	exists.	cow-
lɔŋ	*áh elɔ́ŋ*			
DCM	be big			

'I have a few cows. The cows are big.'

Note that, in the second sentences of (7) and (8), the NPs d͡ʒuŋwɔ́-lɔŋ 'the (many) cranes' and mayén-lɔŋ 'the cows', respectively, serve as the definite counterparts to the indefinite NPs d͡ʒuŋwɔ́ héntékték 'many cranes' and mayén ekʰɔkʰɔ 'a few cows' in the first sentences. This data suggests that the bound morpheme -lɔŋ functions as a definite non-numerical plural quantifier, correlating with héntékték 'many' and ekʰɔkʰɔ 'a few'. Following Borah (2011; forthcoming), we designate -lɔŋ as a definite collective marker (DCM).

It is important to note that no numeral can be used with the suffix -lɔŋ, which means it is not a plural marker like the -s in English. Instead, -lɔŋ indicates a definite collection whose definiteness and imprecise numerosity cannot be overridden (like bor in Assamese). Therefore, it

2 Drawing on Borah's (2013; forthcoming) research, it becomes evident that many Tibeto-Burman languages, along with languages from other families spoken in Northeast India, possess a set of inherently indefinite quantifiers that quantify nouns for an indefinite, unspecified quantity or measure of the noun referent. Corresponding to these indefinite quantifiers is a set of Definite Collective Markers (DCMs) that function as their definite counterparts, quantifying the noun to yield the definite counterpart of that unspecified quantity or measure. While these languages employ such quantifiers for imprecise quantification, they utilize numerals (often with classifiers) for precise quantification, often with distinct definite marking. This broader typological context, as described by Borah, provides a valuable framework for understanding the system of reference in languages of this region.

is more accurately described as a definite collective marker (DCM) rather than a plural marker (see Borah 2008; Footnote 1).

Demonstratives in Tutsa: Referring in physical space

Demonstratives are words like 'this/these' and 'that/these' that act as a direct bridge between language and the physical world. The term itself comes from the Latin dēmōnstrāre, meaning 'to point out,' as the speaker uses them to demonstrate the referent, often with a physical gesture. This is why demonstratives are deictic tools; their meaning depends on the physical context, such as the speaker's location and direction.

By their nature, demonstratives require both the speaker and listener to be physically present and share an awareness of the object being referenced. This object must be visible or otherwise accessible to both parties in their shared space.

Demonstratives are linguistic universals because all languages have a way to refer to objects based on their spatial location. This is tied to our fundamental human need to specify objects in our immediate environment. The concepts of proximal (near the speaker, e.g. 'this') and distant (far from the speaker, e.g. 'that') are basic and universal human experiences that every language encodes into its grammar.

Finally, demonstratives are inherently definite because they presuppose a specific, identifiable referent that is already known to the listener. As linguist Stephen C. Levinson (1983) states in his work on deixis, a definite expression requires that the hearer can identify the intended referent from the context: "The use of a definite expression presupposes that the hearer can identify the intended referent from the linguistic or non-linguistic context." By directly pointing to a specific object, demonstratives fulfill this condition, making them a type of definite expression that does not introduce a new, unspecified item.

It is worth discussing a relevant point about definiteness at this point. Abbott (2019) has thoroughly explored the concept of definiteness, noting that linguists generally take one of two main views: either definiteness means something is unique, or it means something is already familiar. While a sense of familiarity often comes across when we use definite terms, Abbott suggests that the true, underlying meaning of definiteness is actually about uniqueness. She provides examples where something is clearly definite, even if it is not familiar to the listener. Her research offers a vital perspective for examining why, despite all proper names referring to something unique, some might need an explicit 'the' (or a similar marker) and others might not, depending on the specific rules of a language.

From this perspective, the demonstrative serves as the explicit linguistic tool that allows a speaker to enforce uniqueness and achieve successful reference, regardless of whether the hearer was previously familiar with the referent. Even when an object is visible to both the speaker and hearer, it is the speaker who performs the mental act of singling out that specific entity. When a speaker says 'that book,' they are not assuming the hearer's prior knowledge; rather, they use the demonstrative to linguistically enforce the entity's uniqueness, enabling the hearer to identify the intended referent from the context (e.g. a pointing gesture or a shared visual field).

In this scenario, the entity is unique by virtue of the speaker's act of demonstrating it, even if it is not familiar to the hearer beforehand. This is the core of Abbott's argument and the reason her uniqueness-based view is considered a more robust explanation for definiteness than the familiarity-based view.

Our Tutsa data reveals that Tutsa employs two demonstratives: ará ‘this/these’ and erá ‘that/those’, as in (9) and (10) below:

(9) ará ledáp
this book
‘This book.’

(10) erá ledáp
that book
‘That book.’

However, demonstratives are not just for pointing to things that are visible; they also point to concepts in our mental space. This is a crucial function that helps create cohesion in conversation.

In English, for example, when someone says, ‘That’s not true,’ the demonstrative ‘that’ is not pointing to a physical object but to a previously mentioned idea or statement. This same principle applies to Tutsa, as the examples (11) and (12) show.

In example (11), erá ‘that’ does not refer to a visible item but to a previously stated concept. The speaker is referring back to a statement or idea that has been introduced, which exists only in the shared mental space of the conversation. Similarly, in example (12), ará ‘this’ refers to a preceding thought or idea that the speaker considers correct. Both uses demonstrate how demonstratives can serve as anaphoric referring expressions, allowing speakers to connect new information to existing mental referents without having to repeat the full phrase.

(11) erá ta-ə̀ŋka áh
that not true be
‘That is not true.’

(12) ará eláŋe áh
This correct be
‘This is correct.’

By way of another example, the demonstrative ará ‘this’ is used to specify a particular bamboo basket that is physically present and likely visible or within the immediate reach of the speakers. This anchors the referent in the physical context. In contrast, the other demonstrative, erá ‘that’, used in the same context, refers to the same basket, but it identifies it as an object that was talked about in the past before the hearer. Thus, erá refers to a mental concept of the basket, which was formed in the past and is therefore distant. Thus, in (13), the demonstrative erá ‘that’ connects the physically present basket to a specific memory or concept of it in their shared past. It distinguishes between the basket as a current physical object and as a mentally distant conceptual one from a past conversation.

(13) ará tʰèkɔk áh erá tʰèkɔk
this bamboo basket be that bamboo basket
‘This bamboo basket is that basket I talked about earlier.’

Generic Reference in Tutsa

Generic reference is a fundamental type of noun phrase reference that does not point to a specific individual or a particular group of entities. Instead, it refers to a whole class, kind, or category. When speakers use generic reference, their goal is to express generalizations or inherent properties that are considered to be typically true of that category's members. These statements often convey widely held beliefs, scientific principles, or common knowledge about the nature of a particular kind of thing.

In simpler terms, engaging in generic reference allows us to talk about the typical or general features, behaviours, or qualities associated with a whole group, be it animals, objects, concepts, or even abstract ideas, rather than focusing on any single, concrete instance of that group. This ability to make broad statements about categories is crucial for conveying general knowledge, formulating rules, and expressing our understanding of the world in a comprehensive way. While individual members of a category might deviate from these generalizations, generic reference captures the essence or the perceived norm of that kind (see e.g. Leslie and Zoe, 2016: 191-209; Radden 2009: 274-292).

The Bare noun as a generic referential expression

Our analysis of Tutsa data reveals that the language employs a bare noun to express generic reference, as already observed in Example (2) above, reproduced as (14) below. Like (14), Example (15) also uses a bare noun to make a generic statement about an entire class of animals. In example (14), the sentence d͡ʒuŋwɔ áh epùŋ 'The crane is white' does not refer to a specific, individual crane. Instead, it makes a general assertion about the entire species of cranes, stating that a crane is typically white.

Similarly, example (15), pɔk áh elɔ́ŋ 'The elephant is big in size', is not about one specific elephant. It is a statement of a defining characteristic of the entire class of elephants, asserting that elephants are, by their nature, large animals. In other words, both cases use a bare noun to convey a widely accepted truth or an inherent quality of a species, which is the core function of generic reference.

(14) d͡ʒuŋwɔ áh epùŋ
crane be white
'The crane is white.'

(15) pɔk áh elɔ̀ŋ
elephant be big
'The elephant is big in size.'

Noun-lɔŋ as a generic referential expression

Examining our Tutsa corpus, we observe that generic reference can also be achieved by suffixing the dcm - lɔŋ to the noun, as illustrated in (16) and (17) below. This pattern indicates that Tutsa employs a plural form, marked for definiteness, as another means of expressing general truths or characteristics about a class of entities.

(16) d͡ʒuŋwɔ-lɔŋ áh epùŋ
crane-dcm be white
'Cranes are white.'

(17) pɔk-lɔŋ áh elɔ̀ŋ
elephant- dcm be big
'Elephants are big (in size).'

Thus, generic referring expressions are typically formed in two ways in Tutsa: either by using a bare noun or by suffixing the plural marker -lɔŋ to a noun.

Contrasting these two ways of generic reference

When using a bare noun for generic reference, a speaker is talking about the class of entities in its most abstract form, without any specific or definite examples in mind. For instance, the statement 'Cranes are white' refers to the concept of 'cranes' as a whole.

In contrast, when the -lɔŋ suffix is used, the speaker is referring to a set of entities that, while not specific individuals, are conceptually definite. This allows the speaker to represent the entire class by using a mental shortcut: by discussing a definite set of instances (the dcm form), a metonymic link is created to the entire class. This approach is less abstract because it relies on the idea of a group of definite entities to stand for the whole.[3]

Figure 1 below schematically presents our findings.

<table>
<tr><th colspan="2">Bare NPs as referring expressions</th></tr>
<tr><td>Indefinite/Definite
Singular/Plural
(1) wén wàŋ-hala
guest come-PROG
'A guest is coming./ The guest is coming.'
'Some guests are coming./ The guests are coming.'</td><td>Generic
(2) d͡ʒuŋwɔ áh epùŋ
crane be white
'Cranes are white.'</td></tr>
<tr><th colspan="2">Numerically quantified NPs as referring expressions</th></tr>
<tr><td>Indefinite
(unmarked)
(3) ŋa-é minusá wazɔ́m kaptup-tà-ŋ.
I-ERG girl three see-PST-1
'I saw three girls.'</td><td>Definite
(unmarked)
(4) minusá wazɔ́m áh d͡ʒàsemɔ̀ŋ.
girl three be beautiful
'The three girls were beautiful.'</td></tr>
<tr><th colspan="2">Non-numerically quantified NPs as referring expressions</th></tr>
<tr><td>Noun-Indefinite Quantifier
(5) minusá héntékték wàŋ-tahà
girl many come-PRS PRF
'Many girls have come.'</td><td>Noun-DCM
(6) minusá-lɔŋ áh d͡ʒàsemɔ̀ŋ
girl-DCM be beautiful
'The girls are beautiful.'</td></tr>
</table>

[3] As shown in Borah (2008; forthcoming), this is a common pattern observed in many Tibeto-Burman (TB) languages of North East India, as well as in Assamese, an Indo-Aryan (IA) language, plausibly due to their close proximity.

Demonstrative-marked referring expressions	
ará (this/these) (7) *ará nɔ̀k áh elɔ̀ŋ* this house be big *'This house is big.'*	*erá* (that/those) (8) *erá nɔ̀k áh ehìn* that house be small *'That house is small.'*
Generic Reference	
Bare noun (9) *d͡ʒuŋwɔ áh epùŋ* crane is/are white 'Crane(s) is/are white.'	Noun-DCM (10) *d͡ʒuŋwɔ-lɔŋ áh epùŋ* crane-DCM is/are white 'Cranes are white.'

Table 1 *Reference in Tutsa*

Conclusion

This paper has provided an initial exploration into the system of individuative and generic reference in Tutsa, a largely undocumented Tibeto-Burman language of Arunachal Pradesh, India.

Our analysis, grounded in the observed usage of native speakers, reveals that Tutsa possesses a context-driven system of reference. Unlike languages such as English that rely on dedicated determiners to mark NPs, our analysis reveals that the bare noun in Tutsa is a flexible referential expression, capable of expressing both individuative and generic reference depending on the pragmatic context. This contextual dependence is also evident in numerically quantified NPs, where indefiniteness and definiteness are not grammatically marked but are instead derived from the surrounding discourse. On the other hand, the non-numerically quantified NP is either an indefinite NP, marked by an inherently indefinite quantifier like héntékték 'many', or a definite NP, marked by an inherently definite collective marker like -lɔŋ.

The paper also highlights the function of demonstratives in Tutsa. Beyond their primary role of pointing to objects in physical space, demonstratives like ará 'this' and erá 'that' serve as anaphors and refer to concepts in the mental space as well.

Finally, we showed that Tutsa employs two distinct methods for generic reference: the abstract bare noun and the more concrete, metonymic use of the -lɔŋ suffix. This dual system, also observed in other Tibeto-Burman and Indo-Aryan languages of the region, demonstrates a common typological pattern.

In sum, the Tutsa system of reference seems to heavily rely on a speaker's ability to use context and shared cognitive space to convey precise meaning. This study contributes to the broader understanding of how languages in North East India have developed unique and effective strategies for grounding discourse.

While this study offers a foundational understanding of reference in Tutsa, several avenues for future research emerge. Investigating the interplay between these reference strategies and other grammatical features, such as case marking, verb agreement, and discourse particles, would provide a more comprehensive picture of Tutsa grammar. Moreover, examining the use of these reference mechanisms in diverse discourse contexts, including narratives, conversations, and

procedural texts, could reveal pragmatic nuances and information flow strategies within the language.

Finally, comparative studies with other Tibeto-Burman languages of Northeast India would be invaluable for identifying broader typological trends and language-specific innovations in the expression of reference within this rich linguistic area. This initial description contributes to the growing documentation of lesser-studied Tibeto-Burman languages and lays the groundwork for deeper investigations into the intricacies of Tutsa grammar and its place within the linguistic situation of Northeast India.

References

Abbott, B . (2019). Definiteness and familiarity. In Jeanette & and Barbara Abbott (Eds). *The Oxford Handbook of Reference*, (pp 117-129). Oxford: Oxford University Press.

Borah, G. K. (2008). *The Indeterminacy of the Bare Noun and Classifiers: The Case of Assamese.* Unpublished PhD dissertation submitted to Norwegian University of Science and Technology.

Borah, G. K. (2012). Classifiers in Assamese and their meaning chains. In S. Morey, M. Post & G. Hyslop (Eds.), *North East Indian Languages,* (Vol. 4, pp. 292-314). New Delhi: Cambridge University Press.

Borah, G. K. (forthcoming). Noun phrase structure across language families in Northeast India. In Umarani Pappuswamy (Ed.), *Advances in Linguistic and Cultural Studies in Multilingual India: A Commemoration Volume in Memory of Omkar Nath Koul*. Springer.

Das, B. (2021). Auxiliary verbs of Nocte, Khappa, Ollo and Tutsa. *Himalayan Linguistics, 20*(2), *Languages and Peoples of the Eastern Himalayan Region*, (pp 47-60). https://escholarship.org/uc/item/4602d4p6.

Hammarström, H., Forkel, R., Haspelmath, M., & Bank, S. (2025). *Glottolog 5.2*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at https://glottolog.org/resource/languoid/id/tuts1235).

Hanks, P. (2019). Reference as a speech act. In Jeanette Gundel & and Barbara Abbott (Eds.), *The Oxford Handbook of Reference*, (pp 117-129). Oxford: Oxford University Press.

Konch, H. (2021). Nominal Inflection of the Tutsa Language. *International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN:2278-3075 (online), 10*(4), (pp 138-140). https://doi.org/10.35940/ijitee.D8428.0210421 .

Leslie, S-J., & Lerner, Z**.** **(**2016). Generic reference. In A. Capone & F. Lo Piparo (Eds.), *The Handbook of Linguistic Pragmatics* (pp 191-209). Cham, Switzerland: Springer.

Levinson, S. C. (2003). *Pragmatics.* Oxford, England: Oxford University Press.

Lewis, M. P., Simons, G. F., & Fennig, C.D. (Eds.). (2014). *Ethnologue*: *Languages of the World* (17th ed.). SIL International.

Matisoff, J. A., with Baron, S. P., & Lowe, J. B. (1996). *Languages and dialects of Tibeto-Burman* (STEDT Monograph Series 2). Center for Southeast Asia Studies, University of California, Berkeley.

Morey, S. (2017). Tangsa. In G. Thurgood & R. J. LaPolla (Eds.), *The Sino-Tibetan Languages* (2nd ed., pp. 350-368). Routledge.

Radden, G**.** (2009). Generic reference in English: A metonymic and conceptual blending analysis. *Metaphor and Symbol, 24*(4), (pp 274-292).

Rao, N. S. (2010). *The Tutsa Tribe of North East India, Parts 1 and 2*. New Delhi: Mittal Publication.

Rekhung, W. (1992). *Tutsa Language Guide*. Itanagar: The Director of Research, Government of Arunachal Pradesh.

Vivekananda Kendra Institute of Culture. (2005/2016). *Traditional Systems of the Tangsa and the Tutsa*. Guwahati: VKIC.

Abbreviations

DCM	Definite Collective Marker
ERG	Ergative
LOC	Locative
PST	Past
PRS PRF	Present Perfective
PROG	Progressive

Jadavpur Journal of Languages and Linguistics
ISSN: 2581-494X

একক পদ ও একক বাচ্য

পিনাকী সরকার[1]
ভিক্টরিয়া ইন্সটিটিউশন (কলেজ), পশ্চিমবঙ্গ

ARTICLE INFO
Article history:
Received 15/09/2025
Accepted 07/10/2025
Keywords:
বাচ্য,
ব্যক্তি নাম,
একক পদ,
নির্দিষ্ট বর্ণনা

ABSTRACT
ফ্রেগের হাত ধরে বাচ্য ও তাৎপর্য বিষয়ে যে বিতর্কের সঙ্গে আমরা পরিচিত হয়েছি, একক পদ (Singular term) বিষয়েও তা সমান ভাবে প্রাসঙ্গিক। বর্তমান প্রবন্ধে একক পদ বলতে আমরা কোন্ পদগুলিকে বুঝবো তা নিয়ে আলোচনা করা হয়েছে, কেননা কোন্ পদগুলিকে একক পদ বলব তা নিয়ে ভাষাদর্শনে মতভেদ আছে। যেমন- ব্যক্তিনাম বা নির্দিষ্ট বর্ণনাকে সাধারণত একক পদ বলা হলেও, রাসেল এই মতের সঙ্গে সহমত পোষণ করেন না। অতঃপর এই একক পদগুলির অর্থবোধে বাচ্য (Reference) অথবা তাৎপর্যের মধ্যে কোন্‌টি সহায়ক, ফল স্বরূপ কোন্ পদকে একক পদ বলা হবে তার বিচার করা হয়েছে। ফ্রেগে এবং রাসেলের বর্ণনাতত্ত্বের পর, কীভাবে আবার বাচ্যতত্ত্বের পুনর্জন্ম হচ্ছে, একক পদের আলোচনাকে কেন্দ্র করে সেই আলোচনাও প্রাসঙ্গিক হয়ে উঠেছে। পরিশেষে একক পদের বাচ্য বিষয়ে একটি সিদ্ধান্তে আসার চেষ্টা করা হয়েছে যা রাসেলের মতের পরিপন্থী বলেই মনে হবে।

বাচ্য প্রসঙ্গে

ভাষাদর্শনে বাচ্যতত্ত্ব (Theory of Reference) যেভাবে এগিয়েছে, তা আমাদের চিন্তার বিষয়বস্তু কী হবে, তাকে বিশেষভাবে প্রভাবিত করেছে। এক্ষেত্রে যে প্রশ্নগুলি স্বাভাবিকভাবেই এসে পড়ে সেগুলি হল- বাচ্যতত্ত্ব আসলে একটি বিরাট জায়গা, সেখানকার কোন্ বিশেষ অংশটি আমাদের চিন্তাকে প্রভাবিত করেছে? বাচ্য বলতেই বা কী বোঝায়? জেনে রাখা দরকার যে, বাচ্য হল- একটি শব্দ বা বাক্য প্রয়োগ করে আমরা যে বিষয় বা বস্তুটিকে বোঝাতে চাই সেটি। যথা- 'রবীন্দ্রনাথ' এই পদটি ব্যবহার করে আমরা আসলে জগতের যে ব্যক্তিটিকে বোঝাতে চাই সেটিই হল 'রবীন্দ্রনাথ' এই পদটির বাচ্য। এই বিষয়ে যে মূল দার্শনিক প্রশ্ন তা হল- একটি ভাষা বা শব্দ কীভাবে কোনওকিছুকে বোঝাতে পারে? এটি একটি সার্বিক প্রশ্ন এবং ভাষাদর্শনের ক্ষেত্রে এই প্রশ্নটির একটি গুরুত্বপূর্ণ ভূমিকা রয়েছে, কেননা ভাষাদর্শনের একটি অত্যন্ত দৃঢ় ও গুরুত্বপূর্ণ অর্থতত্ত্ব (Theory of Meaning) হল এই যে, ভাষায় উপস্থিত বিভিন্ন শব্দের অর্থ আসলে সেই শব্দটি যাকে নির্দেশিত করে সেই বস্তু বা বিষয়টি অর্থাৎ শব্দটির বাচ্য। তাহলে দেখা যাচ্ছে এইক্ষেত্রে একটি অর্থতত্ত্ব দাঁড়িয়ে রয়েছে তার বাচ্যের উপর। এরপর আরও অনেক জটিল প্রশ্ন রয়েছে, যেমন- আমাদের ভাষার সমস্ত জটিল অংশেরই কি বাচ্য রয়েছে? হসপার্স দেখিয়েছেন যে, একটি বাক্যের সকল অংশের বাচ্য নেই, যেমন- 'এবং', 'অথবা' ইত্যাদির অর্থ থাকলেও বাচ্য নেই। এদেরকে logical connective বা truth functional connective বলা হয়। এদের অর্থ আমরা অনেকভাবে বুঝতে পারি এবং তাদের মধ্যে অন্যতম একটি হল- সত্যসারণীর মাধ্যমে আমরা এদের অর্থ বুঝতে পারি। তাহলে আমরা ভাষায় এমন কতগুলি অংশ বা শব্দ পেলাম যাদের অর্থ বুঝলেও তাদের বাচ্য কী হবে এই নিয়ে আমাদের সমস্যায় পড়তে হল। কাজেই 'একটি পদের অর্থ হল তার বাচ্য' এই মতের বিরুদ্ধে বিভিন্ন যুক্তির সম্মুখীন আমরা হলাম। আবার এমন অনেক নাম

[1] Corresponding Author Email ID: pinaki.sarkar12345@gmail.com

পদের কথাও আমরা বলতে পারি যার অর্থ আমরা বুঝতে পারলেও, তার বাচ্য কী হবে তা নিয়ে নিশ্চিত কিছু বলতে পারি না। যেমন- 'গরু', 'মানুষ' এই পদগুলির অর্থ আমরা বুঝি, কিন্তু তার বাচ্য কী হবে এই নিয়ে সমস্যায় পড়তে হয়। লক্ষণীয় যে, এক্ষেত্রে পদগুলির বাচ্য নেই তা বলা হচ্ছে না, কিন্তু বাচ্য কোন্‌টি তা নিয়ে ধোঁয়াশা রয়েছে। 'গরু' শব্দের দ্বারা কি জগতের সমস্ত গরুকে বোঝাবে? নাকি গোত্ব জাতিকে বোঝাবে? নাকি কোনও বিশেষ গরুকে বোঝাবে? আমরা জানি নামবাদ উপরের কোনও বক্তব্যই মানে না। এছাড়াও আরও এক ধরণের পদ আছে যে পদগুলি যাকে নির্দেশ করে অর্থাৎ তাদের বাচ্য বলতে আমরা যা বুঝি, তার আসলে অস্তিত্বই নেই; তাই সেই পদগুলির বাচ্য আছে- একথা বলা মুস্কিল। যেমন- বিভিন্ন কাল্পনিক চরিত্র বা কাল্পনিক জীব (ফেলুদা, ব্যোমকেশ বক্সী, হাঁসজারু, পক্ষীরাজ ঘোড়া ইত্যাদি)।

বাচ্যতত্ত্ব বিষয়ে এই সমস্যাগুলি আগে থেকেই ছিল এবং আরও নতুন সমস্যা দেখা দিল যখন আমরা ফ্রেগে (Frege, 1948) এবং রাসেলের (Russell, 1905) লেখা পাচ্ছি। এরা দুজনেই অর্থতত্ত্বের আলোচনায় বাচ্যের বদলে যথাক্রমে Sense and Meaning এই দুটির উপর অধিক জোর দিয়েছেন। ফ্রেগে এবং রাসেলকে বলা হয় বর্ণনাত্মক প্রথার (Descriptivist Tradition) ধারক ও প্রচারক অর্থাৎ এরা দুজনেই পদের অর্থবোধের জন্য পদ নির্দেশিত বাচ্যের উপর নির্ভর না করে, সেই পদের বর্ণনার উপর জোর দিয়েছেন। রাসেল ও ফ্রেগের মতের বিস্তারিত আলোচনা করা গেলে পাঠকের পক্ষে সুবিধা হত, তবে পরিসরের অভাবে তা করা সম্ভব হচ্ছে না। এক্ষেত্রে শুধু এটুকু মনে রাখলেই হবে যে, ফ্রেগে ও রাসেল অর্থবোধ বিষয়ে বর্ণনাত্মক প্রথার সমর্থক, কেননা তাঁরা মনে করেন পদের অর্থ বাইরের জগতে সেই বস্তুটি কেমন তার উপর নির্ভর করে না, সংশ্লিষ্ট পদের সঙ্গে যুক্ত বর্ণনাই (ফ্রেগে একে ক্ষেত্রবিশেষে চিন্তাও বলেছেন) পদটির অর্থবোধের সহায়ক।

একক পদ এবং রাসেলীয় বর্ণনাতত্ত্ব

বাচ্য তত্ত্ব ও তার বিরুদ্ধ মতাবলম্বীদের বক্তব্য সম্পর্কে উপরোক্ত পরিচয় দেওয়ার পর আমাদের মূল কাজ হল একক পদ (Singular Term) বলতে আমরা কী বুঝবো তার আলোচনা করা এবং একক পদের ক্ষেত্রে অর্থতত্ত্ব বিষয়ক উপরোক্ত মতবাদ যার নাম বাচ্যতত্ত্ব তা কতদূর সঙ্গত তার আলোচনা করা অর্থাৎ একক বাচ্য (Singular Reference) বলতে কী বোঝায় এবং তা কতদূর সঙ্গত তার পর্যালোচনা করা। ফলত একক পদের বাচ্য সংক্রান্ত আলোচনায়, একক পদ বলতে আমরা কী বুঝবো, তার আলোচনা সঙ্গত কারণেই প্রাসঙ্গিক হয়ে ওঠে। ভাষাদর্শনে মনে করা হয় আমাদের ভাষায় এমন কিছু শব্দ বা শব্দবন্ধ রয়েছে যাদের আমরা একক পদ বলতে পারি এবং এই পদগুলি ব্যবহারের দ্বারা আমরা যে বাচ্যকে বোঝাই সেটিই হল একক বাচ্য। একক পদ বলতে সাধারণত ব্যক্তিনাম (Proper Name) যেমন- রবীন্দ্রনাথ, কলকাতা ইত্যাদি, নির্দিষ্ট বর্ণনা (Definite Description) যেমন- 'গীতাঞ্জলীর রচয়িতা', 'ভারতের বর্তমান রাষ্ট্রপতি' ইত্যাদি, একক সর্বনাম (Singular Pronouns) যেমন- সে, তিনি ইত্যাদি, নির্দেশক (Demonstratives) যেমন- এটা, ওটা ইত্যাদিকে বোঝায়। অর্থাৎ যেসব পদ কেবলমাত্র একজন নির্দিষ্ট ব্যক্তি বা একটিমাত্র নির্দিষ্ট বস্তুকে বোঝায় তাকে একক পদ বলে।

উপরে একক পদ বিষয়ে আমরা যে কথা বললাম বা একক পদের যে উদাহরণগুলি নিলাম সেখানেই রাসেল বিরুদ্ধ মত প্রদান করে বলেন, অনেক পদ রয়েছে যাদের আমরা একক পদ বলে মনে করি কিন্তু তারা আসলে একক পদ নয়, (Evans, 1982) আর যেহেতু সেগুলি একক পদই নয় কাজেই তাদের একক বাচ্যের বিষয়টিও অপ্রাসঙ্গিক হয়ে পড়ে। রাসেল বিশেষকরে ব্যক্তিনাম ও নির্দিষ্ট বর্ণনার কথা উল্লেখ করে বলেন এইগুলি যথার্থ একক পদ নয়। তাঁর মতে যে বাক্যে এই পদগুলির ব্যবহার হয়েছে আমরা যদি সেই বাক্যের গঠনপ্রণালীকে বিচার করি দেখবো যে সেই পদগুলি আসলে একক পদ নয়। রাসেল বাক্যের যৌক্তিক বিশ্লেষণ করে দেখালেন, একটি বাক্য বা বচন যার ভেতর নির্দিষ্ট বর্ণনা রয়েছে সেটি আসলে কোনও একক বাক্য বা একক বচন নয়, সেটি হল সামান্য বচন (General Proposition)। উদাহরণ- "The

present king of France is bald"- এই বাক্যের পূর্বগ (Subject) হল "The present king of France" এবং অনুগ (Predicate) হল "bald"। রাসেল বলেন এই বাক্যের পূর্বগ অংশটিকে আমরা একক পদ ভেবেছিলাম এবং সেই সঙ্গে তাকে বাচ্যার্থক পদও ভেবেছিলাম, কিন্তু তা সঠিক নয়। এই বাক্যের প্রকৃত অর্থ হল- There is X such that X is a king of France and for all Y, if Y is a king of France then Y is the same as X and X is bald. রাসেল বলেন এই বাক্যটি একটি বিশেষ বচন যার পরিধির মধ্যে একটি সামান্য বচন ঢুকে রয়েছে, তাই বচনটি একটি একক বচন হলনা, এটি একটি সামান্য বচন হয়ে গেল। আর যদি বচনটি একটি সামান্য বচন হয় তাহলে তার পূর্বগতে যে পদটি রয়েছে সেটি কোনও একক পদ হতে পারে না এবং তাই তার একক বাচ্যের প্রসঙ্গটিও নিরর্থক হয়ে যায়। রাসেল আরও বলেন যে আমাদের ভাষায় উপস্থিত সকল ব্যক্তিনামই আসলে নির্দিষ্ট বর্ণনার ছদ্মরূপ (Proper names are the definite descriptions in the disguise of a name)। কেননা ব্যক্তি নামের সঙ্গে আমরা কোনও না কোনও নির্দিষ্ট বর্ণনা যুক্ত করি এবং যে ব্যক্তি ওই নির্দিষ্ট বর্ণনার সঙ্গে যুক্ত শর্তকে পূরণ করতে পারে তার ক্ষেত্রেই নামটি প্রয়োগ হয়, অন্য কারো ক্ষেত্রে নয়। অর্থাৎ রাসেলের মতে ব্যক্তিনাম ও নির্দিষ্ট বর্ণনা এই দুটির কোনওটিই একক পদ নয় এবং সেই কারণেই তাদের একক বাচ্যের প্রসঙ্গটিও অনর্থক।

রাসেল মনে করেন "The present king of France" এই নির্দিষ্ট বর্ণনাটির বা কোনও ব্যক্তিনামের বাচ্য নির্দিষ্ট হতে পারে শুধুমাত্র কিছু বর্ণনার দ্বারা। তাহলে এমন একটা নির্দিষ্ট বর্ণনা বা ব্যক্তিনাম থাকতে পারে যার হয়তো কোনও বর্ণনা যুক্ত রয়েছে কিন্তু বাইরের জগতে এমন কোনও কিছুই নেই যার সঙ্গে ওই বর্ণনাটি সঙ্গতিপূর্ণ। কিন্তু তার মানে এই নয় যে ওই নির্দিষ্ট বর্ণনা বা ব্যক্তিনামটি অর্থহীন হয়ে যাবে। অর্থাৎ যদি আমি বলি ব্যোমকেশ বক্সীর অনুমান ও বিশ্লেষণ ক্ষমতা অত্যন্ত তীক্ষ্ণ- এই বাক্যটি অর্থপূর্ণ, যদিও ব্যোমকেশ বক্সী নামে কেউ এই জগতে নেই। অর্থাৎ ব্যোমকেশ বক্সী নামে কেউ এই জগতে না থাকলেও, বাক্যটি অর্থহীন হয়ে পড়ছে না কেননা এই নামের সঙ্গে যুক্ত যে বর্ণনা তা ব্যক্তিনামটিকে এবং সেইসঙ্গে এই নামযুক্ত সম্পূর্ণ বাক্যটিকে অর্থপূর্ণ করেছে। বর্ণনাত্মক প্রথার এটিই মূল বক্তব্য।

এবার রাসেল বিশুদ্ধ একক পদ বা যৌক্তিক একক পদ (Purely Referring Expression or Logical Referring Expression) সম্পর্কে তাঁর মত দিলেন। রাসেলের মতে- একটি বিশুদ্ধ একক পদ হল সেটি যার একমাত্র কাজ হল বাচ্যকে নির্দেশ করা এবং যদি সেই পদ এই কাজটি করতে ব্যর্থ হয়, তবে সেই পদটি অর্থহীন হয়ে পড়ে; ফলস্বরূপ যে বাক্যে পদটি ব্যবহার হয়েছে সেই বাক্যটিও অর্থহীন হয়ে পড়বে। রাসেলের এই শর্ত অনুযায়ী 'ব্যোমকেশ বক্সীর' মতো ব্যক্তিনাম বা 'ভারতের বর্তমান রাষ্ট্রপতি' এর মত নির্দিষ্ট বর্ণনাকে একক পদ বলা যায় না, ফলস্বরূপ তাদের একক বাচ্যের প্রসঙ্গও আসেনা।

এখন প্রশ্ন হল, তবে কারা বিশুদ্ধ একক পদ হবে এবং কাদেরই বা বিশুদ্ধ একক বাচ্য থাকবে? রাসেল বলেন নির্দেশক (Demonstratives) যথা- এটা, ওটা (This, That) ইত্যাদি, যা মানুষের ব্যক্তিগত ইন্দ্রিয় উপাত্তকে নির্দেশ করে। রাসেল এমন বললেন, কেননা তাঁর একটি নীতি রয়েছে যার নাম Acquaintance Principle (Russell, 1912) এবং এই নীতি অনুসারে তিনি বলেন, আমি সেই বচনকেই সম্পূর্ণরূপে জানতে পারি, যার প্রত্যেকটি অংশ আমি Acquaintance এর দ্বারা পেয়েছি। এর থেকেই অনুসরণ করে যে আমি তাকেই বিশুদ্ধভাবে নির্দেশ করতে পারি যাকে আমি Acquaintance এর মাধ্যমে জেনেছি। অর্থাৎ রাসেলের মতে, শুধুমাত্র Acquaintance দ্বারা প্রাপ্ত জ্ঞান আমাদের ব্যক্তিগত ইন্দ্রিয় উপাত্তের মাধ্যমে হয় আর সকল ক্ষেত্রেই আমাদের জ্ঞান হয় বর্ণনার (Description) দ্বারা। কাজেই রাসেলকে অনুসরণ করলে বলতে হয়, শুধুমাত্র নির্দেশক (Demonstratives) ছাড়া অন্য কোনও পদকেই একক পদ বলা যায় না এবং তাদের একক বাচ্য রয়েছে সে কথাও বলা যায় না।

বাচ্য তত্ত্বের নতুন যুগ

এরপর আমরা যেন বর্ণনাত্মক প্রথার (Descriptivist Tradition) এক উল্টো হাওয়া দেখতে পাচ্ছি। অর্থাৎ মনে হচ্ছে যেন, ভাষার অর্থবোধে আমরা বর্ণনাত্মক ব্যাখ্যা না দিয়ে আবার বাচ্যের কথা বলছি। অর্থাৎ বাচ্যতত্ত্ব যেন নতুন করে আবার প্রসঙ্গিক হয়ে উঠলো। এই প্রাসঙ্গে আমাদের তিনজন দার্শনিকের নাম উল্লেখ করতে হয়, তাঁরা হলেন- সোল ক্রিপকে, কেথ ডোনেলান এবং হিলারি পাটনাম।

ক্রিপকে বলেন (Kripke, 1980) ব্যক্তিনাম অবশ্যই একটি বিশুদ্ধ একক পদ এবং ব্যক্তিনাম বিষয়ে বর্ণনাতত্ত্ব সম্পূর্ণরূপে ভুল। একটি ব্যক্তিনাম, যাকে রাসেল আসলে নির্দিষ্ট বর্ণনার ছদ্মরূপ বলে মনে করেন, ক্রিপকে সেই প্রসঙ্গ তুলে তাঁর বক্তব্যের ব্যাখ্যা দিয়েছেন। 'ডোনাল্ড ট্রাম্প' এই ব্যক্তিনামের অর্থ রাসেলের কাছে একটি নির্দিষ্ট বর্ণনা যথা- আমেরিকার রাষ্ট্রপতি যিনি ২০১৬ সালে ভোটে জেতেন। অর্থাৎ যতবার আমি 'ডোনাল্ড ট্রাম্প' এই ব্যক্তিনামের ব্যবহার করছি, রাসেলের মতে, আসলে আমি বলছি আমেরিকার রাষ্ট্রপতি যিনি ২০১৬ সালে ভোটে জেতেন। কিন্তু মনে রাখতে হবে ডোনাল্ড ট্রাম্প সম্পর্কে আমরা অনেক কথা বলতে পারি; যথা- তাঁর জীবনে কী কী ঘটেছে, কী কী ঘটতে পারতো প্রভৃতি অর্থাৎ তাঁর সম্পর্কে সমস্ত সম্ভাবনা নিয়েও আমরা কথা বলতে পারি। কাজেই আমরা বলতেই পারি- ডোনাল্ড ট্রাম্প আমেরিকার রাষ্ট্রপতি নাও হতে পারতেন যিনি ২০১৬ সালে ভোটে জেতেন। এখন ডোনাল্ড ট্রাম্প এই ব্যক্তিনামের অর্থ যদি সর্বদা উপরোক্ত নির্দিষ্ট বর্ণনাই হয়, তাহলে এই বাক্যটি দাঁড়ায়- আমেরিকার রাষ্ট্রপতি যিনি ২০১৬ সালে ভোটে জেতেন তিনি আমেরিকার রাষ্ট্রপতি নাও হতে পারতেন যিনি ২০১৬ সালে ভোটে জেতেন। এই বাক্যটি একটি স্ববিরোধী বাক্যে পরিণত হল কিন্তু আগের বাক্যটি স্ববিরোধী নয়। কাজেই প্রতিটি ক্ষেত্রে 'ডোনাল্ড ট্রাম্প' মানে উপরোক্ত নির্দিষ্ট বর্ণনাই হবে এই কথাটি সত্য নয়। ক্রিপকের মতে যখন আমরা কোনও ব্যক্তিনামের দ্বারা কাউকে বোঝাতে চাই, তখন তাকে আমরা সকল সম্ভাব্য জগতেই বোঝাতে চাই। সকল সম্ভাব্য জগৎ বলতে বুঝতে হবে সকল প্রকার সম্ভাবনা (অর্থাৎ যা হয়েছে, যা হতে পারে বা যা হচ্ছে) নিয়ে কথা বলতে পারা। এই ক্ষেত্রে ওই বর্ণনাগুলি সর্বদা কাজ করেনা। তাই আমাদের এমন একটি উপায়ের কথা ভাবতে হবে যার দ্বারা আমরা কোনও বর্ণনা ব্যতিরেকেই ওই ব্যক্তিকে বোঝাতে পারবো। সেই কারণেই ক্রিপকে, ব্যক্তিনামকে বলেছেন- Rigid Designators অর্থাৎ যা দৃঢ়ভাবে নামকরণ করে।

কেথ ডোনেলান রাসেলের বিরোধিতা করে বলেন (Donnellan, 1966) একটি নির্দিষ্ট বর্ণনারও বিশুদ্ধ বাচ্যিক ব্যবহার হতে পারে। তিনি একটি কল্প পরীক্ষার সাহায্য নিয়ে বলেন, ধরাযাক একজন ব্যক্তি যার নাম স্মিথ, সে খুন হয়েছে এবং তাকে এমন নৃশংসভাবে খুন করা হয়েছে যে সকলেই বলছে- স্মিথের খুনী একজন উন্মাদ ব্যক্তি। আরও ধরাযাক, পুলিশ একজন আসামিকে ধরেছে এবং দাবি করছে সেই ব্যক্তি স্মিথকে খুন করেছে এবং তাকে আদালতে আনা হয়েছে। আদালতে দাঁড়িয়ে সেই আসামি এমন সব বিচিত্র অঙ্গভঙ্গি করছে যে তাকে দেখে কেউ কেউ বলছে- স্মিথের খুনী একজন উন্মাদ ব্যক্তি। এই ক্ষেত্রে একটি বর্ণনা একটি বিশুদ্ধ বাচিক পদ হিসেবে কাজ করছে। কেননা এমন হতেই পারে যে, হয়তো ওই ব্যক্তি স্মিথের খুনী নয়, কিন্তু উক্ত বর্ণনা (স্মিথের খুনী) প্রয়োগ করে আমরা সেই ব্যক্তিকেই নির্দেশ করছি, কেননা ওই ব্যক্তিই স্মিথের খুনী হিসেবে সাব্যস্ত হয়েছে। যদিও এই বর্ণনা তার উপর খাটে না তবুও এই বর্ণনা প্রয়োগ করে আমরা সেই ব্যক্তিকে নির্দেশ করছি। হতেই পারে, যে আসলে খুন করেছে সে নিজেই ওই ব্যক্তির উদ্দেশ্যে উক্ত বর্ণনা প্রয়োগ করছে এবং তার বাচ্য বা নির্দেশনা সফলও হচ্ছে। অর্থাৎ এই ক্ষেত্রে আমরা দেখলাম যে একটি নির্দিষ্ট বর্ণনারও বিশুদ্ধ বাচ্যিক ব্যবহার হতে পারে। অর্থাৎ একটি নির্দিষ্ট বর্ণনার একটি নামের মতো ব্যবহার হচ্ছে, আরও বলা যায় যে, একটি নির্দিষ্ট বর্ণনার বাচ্যের প্রসঙ্গও উঠে আসছে, যদিও রাসেল মনে করেন একটি নির্দিষ্ট বর্ণনার বাচ্যের প্রসঙ্গ ওঠে না।

আমরা এতক্ষণ যে আলোচনা করলাম তা মুলত একক পদ নিয়ে, কেননা এটা মনে করা হয় যে একটি সার্বিক পদের (General Term) বাচ্য না থাকলেও একটি একক পদের বাচ্য থাকার সম্ভাবনা প্রবল। কিন্তু এখানে আমরা হিলারি পাটনামের একটি বিখ্যাত লেখা (Putnam, 1997) পাই যেখানে তিনি মুলত

সার্বিক পদগুলির বাচ্যের কথাই তুলে ধরেন। তিনি তাঁর বিখ্যাত যমজ জগত কল্পপরীক্ষার (Twin Earth Thought Experiment) মাধ্যমে দেখান যে, একটি পদের অর্থকে বুঝতে গেলে সেই পদটি বাইরের জগতে যে বস্তুকে নির্দেশ করে তার প্রেক্ষিতেই বুঝতে হবে। যথা- জল এই পদের অর্থকে বুঝতে গেলে বাইরের জগতে আমরা যে তৃষ্ণা নিবারণকারি তরল পদার্থটি নিয়ে নাড়াচাড়া করি, তার মাধ্যমেই বুঝতে হবে। নিশ্চিতরূপে এই বক্তব্য বাচ্যতত্ত্বের সমর্থনে কথা বলে।

কাজেই ক্রিপকে, ডোনেলান এবং পাটনামের হাত ধরে আমরা যেন নতুন করে আবার পদের অর্থ বোধে বাচ্যতত্ত্বকেই প্রাধান্য দিতে শুরু করলাম। আর বাচ্যতত্ত্বকে প্রাধান্য দেওয়ার অর্থই হল একক পদের সমর্থনে কথা বলা যাদের একক বাচ্য রয়েছে। আর পাটনামকে অনুসরণ করে বলা যায় যে শুধুমাত্র একক পদ নয়, সামান্য পদগুলিরও বাচ্য রয়েছে এবং তাদের বাচ্য কোনও বর্ণনার উপর নির্ভর করেনা বরং নির্ভর করে জগতে সেই বস্তুটি কিরকমভাবে রয়েছে তার উপর। কাজেই একক পদ ও তাদের একক বাচ্যের সমর্থনে আমরা যুক্তি পেলাম।

একক পদ ও তার বাচ্য বিষয়ে সিদ্ধান্ত

ভাষায় যে একক পদ রয়েছে তা নিয়ে মতভেদ নেই কিন্তু কোন্ পদ গুলিকে একক পদ বলা হবে এবং এই পদগুলির অর্থবোধে বাচ্য না তাৎপর্য, কোন্‌টি কাজে লাগে সেটি নিয়েই আসল বিতর্ক। আরেকটি বিতর্ক হল ভাষার কোন্ কোন্ প্রকারের পদকে একক পদ বলব তার তালিকা নিয়ে। আমরা দেখলাম সাধারণভাবে ব্যক্তিনাম এবং নির্দিষ্ট বর্ণনাকে একক পদ মনে করা হলেও, রাসেল তাঁর যুক্তির সাহায্যে দেখাতে চাইলেন আসলে নির্দিষ্ট বর্ণনা একক পদ নয়। সেইসঙ্গে প্রতিটি ব্যক্তিনাম আসলে নির্দিষ্ট বর্ণনার ছদ্মরূপ হওয়ায় ব্যক্তিনামও একক পদ নয়। রাসেলের আলোচনা থেকে এটি বেড়িয়ে আসে যে, একক পদের অর্থবোধে বাচ্য মূল ভূমিকা পালন করে। কেননা রাসেলের মতে একক পদের একমাত্র কাজই হল বাচ্যকে নির্দেশ করা, যদি একটি একক পদ বাচ্যকে নির্দেশ করতে না পারে তাহলে সেই পদ অর্থহীন হয়ে পড়ে।

রাসেল যে প্রচলিত একক পদগুলিকে প্রকৃত একক পদ বলতে চান না, তাঁর এই মতের সীমাবদ্ধতা রয়েছে। তিনি বলেন, 'রবীন্দ্রনাথ', 'ভারতের বর্তমান রাষ্ট্রপতি', এমনকি 'ব্যোমকেশ বক্সী'- এদের আমরা একক পদ ভাবলেও, আসলে তারা নির্দিষ্ট বর্ণনার ছদ্মরূপ মাত্র। কিন্তু এতে ভাষার ব্যবহারিক ও সামাজিক বাস্তবতাকে উপেক্ষা করা হয়। প্রতিদিনের ভাষা ব্যবহার ও সাহিত্যিক প্রকাশে ব্যক্তিনামকে একক পদ হিসেবেই ব্যবহার করা হয়। রাসেল যে ধরণের যৌক্তিক বিশ্লেষণ দেখান, সেই যুক্তি সাধারণ ভাষা ব্যবহার থেকে অনেক দূরে অবস্থিত। অন্যদিকে রাসেল বলেন, অস্তিত্বহীন সত্তা নিয়েও গঠিত বাক্য অর্থপূর্ণ হতে পারে (যেমন, ব্যোমকেশ বক্সী নেই, তবু বাক্যটি অর্থপূর্ণ)। কিন্তু আবার বলেন, যদি একটি পদ একক পদ হয় এবং তা কাউকে নির্দেশ করতে না পারে, তবে সেটি অর্থহীন হয়ে যাবে। এখানে দ্বৈততা রয়েছে। একদিকে, অস্তিত্বহীন পদকে নিয়েও অর্থপূর্ণ বাক্য সম্ভব বলছেন, আবার অন্যদিকে একক পদের বাচ্য না থাকলে বাক্য অর্থহীন হয়ে যাবে বলছেন।

রাসেলের Acquaintance Principle অনুযায়ী, আমি কেবল সেই বস্তুকেই ঠিকভাবে উল্লেখ করতে পারি অর্থাৎ সেই বস্তুরই বাচ্যের কথা বলেতে পারি যাকে আমি প্রত্যক্ষভাবে চিনি বা জেনেছি। কিন্তু এক্ষেত্রে প্রশ্ন ওঠে, ভাষা কি কেবল ব্যক্তিগত ইন্দ্রিয় অভিজ্ঞতায় সীমাবদ্ধ? ইতিহাস, বিজ্ঞান বা সাহিত্য সংক্রান্ত বহু বিষয় আমরা সরাসরি জানি না, তবুও আমরা তাদের নিয়ে অর্থপূর্ণ বাক্য রচনা করি বা প্রয়োগ করি। আমরা যখন বলি, "নেপোলিয়ন ওয়াটারলুর যুদ্ধে পরাজিত হন।" – রাসেলীয় যুক্তিতে এই বাক্যটি অবান্তর হয়ে পড়বে বলেই মনে হয়, কেননা আমি নেপোলিয়নকে ব্যক্তিগতভাবে চিনি না। প্রণব কুমার সেন রাসেলের acquaintance এর ধারণার অসঙ্গতি প্রসঙ্গে বলেন, এই ধারণা সঠিক হতে পারে না (Sen, 1991)। আমরা যাদের proper name অর্থাৎ ব্যক্তিনাম নামের মাধ্যমে উল্লেখ করি, তাদের মধ্যে কতজনের সঙ্গে আমরা পরিচিত? আমরা কি বিশ্বের সব শহর বা ঐতিহাসিক ব্যক্তিত্বদের সঙ্গে পরিচিত? সুতরাং,

কেবলমাত্র acquaintance একজন ব্যক্তির পরিচয় জানার একমাত্র উপায় হতে পারে না অর্থাৎ, একজন নির্দিষ্ট ব্যক্তি কে তা জানার জন্য শুধু এই উপায় যথেষ্ট নয়। এই জানার অন্যান্য উপায় থাকতে পারে এবং এই অন্যান্য উপায়গুলো এমন, যেগুলোর ওপর দার্শনিকেরা হয়তো বিশেষ গুরুত্ব দেননি। জানার এই উপায় হতে পারে- সামাজিক ও সাংস্কৃতিক জ্ঞান বা ঐতিহ্যের মাধ্যমে জানা (Sen, 1991)।

উপরের আলোচনার পর একক নাম বিষয়ে আমাদের যদি একটি সিদ্ধান্তে পৌঁছোতে হয়, তাহলে আমি ব্যক্তিনাম, নির্দিষ্ট বর্ণনা এবং নির্দেশক এই তিন প্রকারের পদকেই একক পদের মর্যাদা দেওয়ার পক্ষপাতি। এক্ষেত্রে রাসেলের মতের সীমাবদ্ধতার কথা উল্লেখের পাশাপাশি কেন এই তিন প্রকারের পদকেই একক পদ বলব, সেই বিষয়ে যুক্তি দেওয়া প্রয়োজন। একক পদ বিষয়ে যে সবচেয়ে কথাটি আমাদের মনে রাখতে হবে তা হল, পদটি ব্যবহার করে আমরা কেবলমাত্র একজন ব্যক্তি বা বস্তুকেই বোঝাই, একাধিক ব্যক্তি বা বস্তুকে নয়। একক পদের যে আবশ্যিকভাবে বাহ্য জগতে একটি বাচ্য থাকতেই হবে তেমন নয় (যেমন আমরা আগেই দেখলাম, কাল্পনিক নাম যথা 'ব্যোমকেশ বক্সী'র কোনও নির্দিষ্ট বাচ্য নেই)। বরঞ্চ আমাদের দেখতে হবে, যদি সেই পদটির কোনও বাচ্যের কথা আমরা মনে মনে ভাবি বা চিন্তা করি, তাহলে সেই পদটির একাধিক বাচ্য আছে, এই কথাটি নির্দ্বিধায় ভাবতে পারছি কিনা অর্থাৎ এমন চিন্তা করতে পারছি কিনা যে, একটি বিশেষ পদের একাধিক বাচ্য রয়েছে (কোনও পদের বাচ্যের কথা মনে মনে ভাবার ব্যাপারে আমাদের সাহায্য করবে যে চিন্তা, ফ্রেগে তাকে sense অর্থাৎ তাৎপর্য বলেছেন)। যদি চিন্তাতেও কোনও একটি পদের একাধিক বাচ্য আছে এটি প্রতিষ্ঠা করতে না পারি, অর্থাৎ এতে চিন্তায় বিরোধিতা দেখা দেয়, তাহলে সেই পদটিকে অবশ্যই একক পদের মর্যাদা দিতে হবে। উদাহরণ দিয়ে বোঝানোর চেষ্টা করি- আমরা সকলেই জানি ব্যোমকেশ বক্সী একটি কাল্পনক চরিত্র যার কোনও বাস্তব বাচ্য নেই। কিন্তু 'ব্যোমকেশ বক্সী' এই নামটির বাচ্যের কথা যদি আমরা চিন্তা করি, তাহলে আমাদের কেউই কি এটা চিন্তা করতে পারবো যে, জগতে একের বেশি বোমকেশ বক্সী রয়েছে? ঠিক যেমন 'গরু' পদের দ্বারা জগতে একাধিক বাচ্যের কথা আমরা ভাবতে পারি বা দেখতে পারি (যথা আমার বারিতে থাকা লাল গরুটি, আমেরিকার কোনও গ্রামে থকা একটি গরু ইত্যাদি) তেমন 'ব্যোমকেশ বক্সী' এই পদের দ্বারা একাধিক বাচ্যের কথা আমরা চিন্তাও করতে পারি না। লেখক তার সৃষ্ট চরিত্র ব্যোমকেশ বক্সীর যে বাহ্যিক ও চারিত্রিক বর্ণনা দিয়েছেন, তার আবশ্যিকভাবে একটিই মাত্র বাচ্য থাকার সম্ভবনা রয়েছে, তার বেশি কখনোই নয়। অর্থাৎ এটা পরিষ্কার যে, এই মানদণ্ডে বিচার করলে নির্দেশকের পাশাপাশি ব্যক্তিনাম ও নির্দিষ্ট বর্ণনাকেও একক পদ বলে স্বীকার করতে হয়। আর কোনও পদের নির্দিষ্ট বাচ্য না থাকেলই যে সেই পদটি বা যে বাক্যে পদটি ব্যবহৃত হয়েছে সেই বাক্যটি অর্থহীন হয়ে পড়েনা, সেটি তো ফ্রেগে ও রাসেল আগেই যুক্তি দিয়ে দেখিয়েছেন। অর্থাৎ প্রথমত, একটি পদ যদি কেবলমাত্র একটি বাহ্য ও বাস্তব বিষয়কে বোঝায় বা উল্লেখ করে, তাকে একক পদ বলব। দ্বিতীয়ত, যদি কোনও পদ বাহ্য ও বাস্তব বিষয়কে বোঝাতে বা উল্লেখ করতে সক্ষম নাও হয়, কিন্তু পদটির দ্বারা কেবলমাত্র একটিই বাহ্য ও বাস্তব বিষয়কে বোঝানোর বা উল্লেখ করা সম্ভব হবে বলে চিন্তা করা যায়, একাধিক বাচ্যের কথা ভাবলে চিন্তায় বিরোধিতা দেখা দেয়, তবে সেই পদটিকেও একক পদ বলতে হবে।

এবার আসি সেই প্রশ্নে যেখানে জিজ্ঞেস করা হয়, একক পদগুলির অর্থবোধে বাচ্য অথবা তাৎপর্যের মধ্যে কোন্‌টি সহায়ক। উপরের আলোচনার ভিত্তিতে আমার উত্তর দুটোই। অর্থাৎ কিছু একক পদের অর্থ বুঝতে বাচ্যের সাহায্য নিতে হয়, কিছু পদের জন্য তাৎপর্যের।

গ্রন্থপঞ্জী

Frege, Gottlob. (1948) Sense and Reference, The Philosophical Review, Vol. 57, Issue. 3.
Russell, Bertrand. (1905). On Denoting, Mind, New Series, Vol. 14, No. 56.
Evans, Gareth. (1982). The Varieties of Reference, ed. J. McDowell, New York: Oxford University Press.

Russell, Bertrand (1912). The Problems of Philosophy, London: Thornton Butterworth Ltd.
Kripke, Saul A. (1980). Naming and Necessity, Cambridge: Harvard University Press.
Donnellan, Keith. S. (1966) Reference and Definite Description, The Philosophical Review, Vol-75, No- 3.
Putnam, Hilary. (1997), The Meaning of Meaning Language and Reality, Mind Language and Reality, Volume 2, Cambridge: Cambridge University Press.
Sen, P. K. (1991). Some Notes on Singular Reference, Reference and Truth, New Delhi: Indian Council of Philosophical Research.

Jadavpur Journal of Languages and Linguistics
ISSN: 2581-494X

Reflection on Linguistic Reality Through Bhartṛhari's Vākyapadīya

Shubham Kumar[1]
University of Delhi, Delhi

ARTICLE INFO

Article history:
Received 15/09/2025
Accepted 06/10/2025

Keywords:
Vedas,
linguistics,
Sūryamāyā,
Bhartṛhari,
Śabdabrahma,
Vākyapadīya.

ABSTRACT

This article raises a problem over the variety of language, specifically in Indian philosophy's grammarian thought and other philosophical doctrines as well. A central query that shall be explored is whether it has any real basis or not, in addition to how language forms concepts. Considered Indian canon when parlance in tandem eternal truth; likewise, scientific unfoldment of natural phenomena in physics—the approach delves into verifying linguistic reality at both transcendental and customary levels. Wherein thinkers widely acclaim the function of words (sound) in deciphering a realm to comprehend the inexplicable, leading to one's extra-lingual-harmonic knowledge.

Introduction

The Collins-English wordbook defines language as the use of a system of communication that consists of a set of sounds or written symbols (Cyberspace, 2024). It is a medium to express one's emotional state, needs, and intellectual understanding. Moreover, to identify the object of speculation, either physically perceivable or having subjective awareness too, pivotal is language's significance.

Kālidāsa opines in the preface of Raghuvaṃśam about intention to word and meaning, tendering their eternity of form along with feature conjoinedly—to realize its essence in their infinitude; he admires Lord Śiva-and-indivisible-Pārvatī (better half), creating the visible world.

vāgarthāviva saṁpṛktau vāgarthapratipattaye/
jagataḥ pitarau vande pārvatīparameśvarau// (1.1).

The conundrum of the universe's origin from śabda or logos, the ultimate and absolute creative principle, was well-known from a very early stage of philosophical speculation in India (Bhattacharya 1985, 2). At the same time, it is corroborated in an aphorism of Rāmānuja's Vedāntasāraḥ—a commentary on Brahma-sūtra refuting the opposing views and quotes.

"śabda iti cennātaḥ prabhavātpratyakṣānumānābhyām" (1.3.27).

Through perception and inference—the mental or spiritual ray—word alone cannot intend the proper status of an object that extends to its enlightenment or definition. For example, a

[1] Corresponding Author Email ID: shubhamkumarbgp22@gmail.com

probandum (sādhya) having luster could only be identified and uttered as 'fire' in hotness; similarly, the inferential disclosure of an entity with the prevalent locutions Īśvara, Oṃkāra, or Saccidānanda. Provided a vocable can never have an import, then the idea could substantially be referred to; otherwise, the purpose of this mundane world is allied to having no objective sense.
Correspondingly, the utmost importance of syllables is found in the following verse to be congruent.

duṣṭaḥ śabdaḥ svarato varṇato vā mithyāprayukto na tamarthamāha/
sa vāgvajro yajamānaṁ hinasti yathendraśatruḥ svarato'parādhāt// (Taittirīya Saṁhitā, 2.4.12).

Elucidation
A word lacking proper accentuation and syllabic pattern is not applicable to a desirable convention for the utterance since a wrong accent put over the reciting 'the enemy of Indra' here acts unfavorably on a speaker.

Rationale
An attempt enumerates the discussed hypothesis of this essay, "Is language real?" while commenting on the text, Vākyapadīya of Bhartṛhari, who roughly comes from the monism (Advaita) school of philosophical classification other than dualistic-pluralistic discourses. The manual formalized by him equates ontological, metaphysical, and epistemological understanding to a word (śabda). As a consequence of that inquiry, word-absolute-being (Śabda-Brahman) observes fascinating ways to allocate the novel approach to propounding.
Sage contends with a philosophy of language through the following verses, some of which are in sequence:

anādi-nidhanaṁ brahma śabdatattvaṁ yadakṣaram/
vivartate'rthabhāvena prakriyā jagato yataḥ// (Vākyapadīya Brahmakāṇḍa, 1).

Brahma, i.e., nontemporal and indestructible, has a rhetorical disposition of letter—the world or creation—a superimposing phenomenon conceptual.

Sūryamāyā
Attaining Brahman is liberation. This is the summum bonum of the goals of humankind. In their effervescent intent, the Veda (right knowledge) leads them where grammar incorporates primary aspects. Differentiation between a sentence and term brings clarity to their exact character; the Vākyapadīya particularly falls into the line of grammatical studies. For this reason, we locate the beginning of the text with a prologue to Brahman's attributes. That all pervading, unifying reality is without other and indivisible further truth, a subject to knowledge wherein distinguishable entities cognate space, time, and objects subdue into un-discriminant.

Changing or illusionary otherwise-abstraction in the form of the world's creative effect is just an imposition; Brahman persists since cause of this: "Yato vā imāni bhūtāni jāyante (3.1.3)," interprets the Taittirīya upaniṣad. Other false awarenesses are similar to snake-rope and dream occurrences. Moreover, every constituent consciousness is identical to that of an eternal, omnipresent, knowing-being and has the scope of linguistic analysis. Furthermore, the cause-and-effect relationship of Brahman to this world is fathomable—closely the subsisting cause (clay) of a pot. That which is the material cause of creating, sustaining, and destroying the universe enunciates Śruti. Here, the objective is to describe two kinds of words: subtle and gross. The presence of the innermost knowing-utterable, i.e., jñāna, in each individual being is a subtle

one, hard to cognize by others in this manner. When there is an amalgamation of letters in any alphabet under the influence of vital force (prāṇa), only a person becomes the medium of knowledge. This is a fabricated (gross) form of the word. Previous is in a slumber, unswerving, inexplicable state, contrary to later, where it gets into falsity on speech content, sentence variation, and its meaning.

In order to remain unshaken in its spirit and appear through alteration, the ground of forming this world principle is 'vivartopādānatva' (projected) of that real (Brahman). Explicated through analyzing the property of some mirror, which could reflect the color (blue or red) and even retain the pigments that fall to it; clean in actuality, picking various colors, merely projecting as such. Whereupon, the Brahman is not an object but rather an omniscient (sarvātmā) or constituting entity (sarvasvarūpa) for the entities—a classification of something perceivable with another resembling or different sense-data-perception is an object, e.g., a tree, which has all three characteristics of being an object, including space and time constraints. Methodology acknowledges rationalizing what could be seen (thinkable) to ponder alongside the unseen (pervasive).

Distinguishable features in Brahman (real) receive their seed from the power of nescience (avidyā). It is neither different nor identical to that. Presuming non-identical, then fail to show its strength, whereas an unreal never be a cause to anything indeed not experienced. Contemplating alike with Brahman, then that shall not be enforced on real-established (sat-siddha); however, it stays obstructed in valid cognition (jñāna) broadly asserted. Hence, the avidyā-śaktis (illusory potencies), unable to determine real or unreal in explicating from both standpoints, advocated mithyā (false) over a peculiar, inexplicable epistemology.

It is intriguing to learn how Brahman holds its creative energy in an onward stanza.

ekameva yadāmnātaṁ bhinnaṁ śaktivyapāśrayāt/
apṛthaktve'pi śaktibhyaḥ pṛthaktveneva vartate// (Vākyapadīya Brahmakāṇḍa, 2).

One-uncountable, although seen in multitude—a substratum of potent; certainly unseparated, not even separable from that—uphold Śrutis (Vedic apprehending).

Illustration: The above verse could have an analogy (udāharaṇa) of a lotus happening in water. However, surviving in it, the flower dwells untouched.

Sūryamāyā

"sadeva somyedamagra āsīdekamevādvitīyaṁ brahma"; O! somya (relevant listener): This world before its creation was that whole, real-like Brahman, unseparated. "praṇava evaikastridhā'bhivyajyate"; The Vedas, Ṛk-Yajuḥ-Sāma, sing the manifesting glory of none other than the syllable Oṃ (praṇava), or else Śabdabrahma.

It is inevitable to doubt—during the time that Brahman's potential is varied, the Āśraya-Brahma (sustaining locus) should be multiple—how is it not so? For example, illuminating power relates to light or photons, which is contradictory to sheathing capability in darkness, where the related things are different, not similar. Answering this question by looking at the peculiar effect, its causal potency could be specific, but the sheltered material cause (upādāna kāraṇa) is not necessarily dissimilar; this is a law (niyama). In the instance of clay, the cause of various earthly products manufactured by a potter in like manner might be multifarious; the material cause of that is doubtless Brahma. Therefore, the existent and its creating force are

wrongly imagined as having a mutually-dependent relationship to their particular cause-and-effect application, whereas all take refuge in it (Brahman), the principle cause.
Preceding verses (ślokas) decipher Bhartṛhari's grammatical mind in philosophizing the nature of words when interpreted to the existence of supreme reality. A similar quest asserted in the Chāndogya upaniṣad, "tadaikṣata bahu syāṃ prajāyeyeti (6.2.3)," expounded that it wants to be the many in form—this is the desire; fulfilling along the dual aspect of śakti (energy) happens to be sorted out in the third aphorism.

avyāhatāḥ kalā yasya kālaśaktimupāśritāḥ/
janmādayo vikārāḥ ṣaḍ bhāvabhedasya yonayaḥ// (Vākyapadīya Brahmakāṇḍa, 3).

Inherent arts or expressions of this (Brahman) are produced through the current of time, since birth or creation, etc., among the six differentiated forms of disturbances that subsist on material causes.

Sūryamāyā
The Brahman's strength to give birth depends on kālaśakti (spatial faculty), which causes six perversions of padārtha (category); these are birth (janma), existence (sat), growth (vardhate), transformation (vipariṇamate), decreasing (apakṣīyate), and destruction (vinaśyati). Knowledge (Veda) is an expedient means to construe this, known as a foremost affirmative. However, naturally, Śabdabrahma has been described as single, with its eternal, indistinguishable material authority, divided birth, and obligatory-time influences. These two do not vary factually; moreover, the formation changes equivalent birth, and the other five get attributed to causal propensity; for that reason, they are called numerous.

Birth has its cause, i.e., time potential; in addition to vice versa, which later pervaded the former in showing multiplicity in spite of the unified singular and, respectively, its dependence on efficient cause (nimittakāraṇa), Brahma, bears the same resemblance. The time mechanism, we usually believe, persists in three states: the past, present, and future. Opposed to the genuineness of the present, another solely lies in our impressions or thought speculation; nevertheless, the momentariness it carries with the existing moment is only actual.

Creation (birth)-existence-destruction, a time force works accordingly in proper control over each, therein marking to know one object in its three deformations anticipating the diversified world. The same water element takes the form of 'bubble-wave-whirlpool as well as glacier-vapor-and-snow' bound to have this mechanism.

Material cause stays inert till the period of the obligatory uniting in purview of beings' (jīvas)' previous-karmic-residue (prārabdha) as an accessory stimulant in initiating action to becoming, not otherwise. The above stanza sets forth a discursive outlook about the beings' appearance based on their accomplishment of deeds in past existences; subsequently, the essence of that (Brahman) is again questionable in its origin:

ekasya sarvabījasya yasya ceyamanekadhā/
bhoktṛbhoktavyarūpeṇa bhogarūpeṇa ca sthitiḥ// (Vākyapadīya Brahmakāṇḍa, 4).

A seed of all is the non-dual graspable multifarious, existing in three categories: sensible beings, sense-enjoyable, and sense enjoyment.

Sūryamāyā
Illustration of the same cognizance-awareness in dream during sleep tends to transform, chariot-horse-charioteer-beside-elephant, seeing empirical images in the Vedānta gospel.
Undefined (neither contrasting nor comparable) time and birth means of that one and only Brahman, unperceivable culminate into beings: "self-mine," etc., egoistic-sensible-being (bhoktā), objects-of-sense-perception (bhogya), homologous sound or touch, in other words, sense-object-contact (bhogakriyā). This differentiation is mutually deviated, whereupon imaginary and false, although the ultimate real (Brahman) happens in illusion but stays pure and illuminating all-together undeviating, unpassable, and unequaled. So, considering the truth, accordingly, a rope in ignorance has an apparent likeness to a snake, water stream, or stick that is not unsteady, grounded in reality (sat), and interdependent with all, for the period that the rope is only existent. Compared with the fact that Brahman's virtue is not inconsistent with it because the existence of both aligns one with the other, not even analogous; corresponding to a nightmare's vision mentioned overhead, this becomes unreal or meaningfully dead. Necessarily, an illusion (māyā) is persuasively undefinable in traits.
The upcoming aphorism of the monograph (kārikā) disclaims ways to attain it (Brahman).

prāptyupāyo'nukāraśca tasya vedo maharṣibhiḥ/
eko'pyanekavartmeva samāmnātaḥ pṛthak pṛthak// (Vākyapadīya Brahmakāṇḍa, 5).

Assimilation into the supreme reality is admitted by great ācāryas, notwithstanding many Vedic preachings that have one parallel understanding.

Sūryamāyā
Ṛk-Yajuḥ-Sāma is naming Śākala-Vākala-Kauthuma et al. streams of cogent discernment (samyak jñāna), which speak for the disciple's deficiencies in enlightening them. The previous stanza corresponds to this in intending with two interchangeable denotations of the absolute soul, yasya (yat) and tasya (tat)—apart howsoever; for example, Praṇava signifies the Brahma or else Viṣṇu-Śiva, an idol of devas.

Waking or dream narratives that are witnessed have been described to others solely through suggested language when a dialect is an expression of someone's inner feeling, akin to a script of letters; whatever audible or inter-actual character implies the Brāhmī-Devanāgarī and so forth. In a similar fashion, aphorism seeker (mantradraṣṭā) sages in deep, flawless insight reveal through resplendent-arising (jñāna) of the subtle, eternally acknowledged-light-in-itself, Śabdabrahma in realization; mantras symbolize that understanding in their complete absorption. It illuminates those unaware followers of the preceptor to meditate on the aforesaid.

Herein, grasping Brahman means the demolition of the ego, i.e., unconscious impressions into the ultimate being (Brahman-like[2])—salvation simultaneously. On the contrary, intellect and individual (I)-consciousness'-recognition over ego's diversification since a false accusation; doer-enjoyer tend worldly bondage.

Now, Bhartṛhari talks about the constitution of the paths that usher in liberating the wise audience.

[2] brahma veda brahmaiva bhavanti (muṇḍakopaniṣad, 3.2.9).

bhedānāṃ bahumārgatvaṃ karmaṇyekatra cāṅgatā/
śabdānāṃ yataśaktitvaṃ tasya śākhāsu dṛśyate// (Vākyapadīya Brahmakāṇḍa, 6).

Multiple dimensions of differences happen to be the limbs of performance only, though the power inherent in words varies amongst its branches.

Sūryamāyā: There lies a special enterprise (karma) evident in miscellaneous Vedic's mode bifurcation; thus, its presence necessitates through all. While reaching one destined place, we discover several pathways in addition to the medium of communication. Furthermore, to gain health, members of Āyurveda are: śalya, śālākya, kāma, bhūta, kaumārabhṛtya, agadatantra, vājīkaraṇatantra, and rasāyanatantra, with characteristic features. Consequently, the didactic mantra enjoined in the variegating knowledge of stanza 'five' results in an implication. Practicing dos, beside abstaining from don'ts, accounts for heaven (svarga), wherefore an end has been anticipated.

In accordance with the Mīmaṃsā philosophy, all branchlets ascertain a solitary achievable exertion—'sarvaśākhāpratyayamekaṃ karma,' properly accepted. How do we explain this equitable dichotomy? Derivation of word has a rule of certain-accents (subtlest form) along-with-breakdown (lopa-āgama-ādeśa) of an alphabet (varṇa) feasible in merit due to appropriate use and meaningful-conceptual ability in arrangement reasons. viz., 'simasyātharvaṇe'nta udāttaḥ; here, Pāṇinī's grammatical system draws a correct exercise of the term "sima" in an ātharvaṇa song ending with a sound (dhvani) 'akāra', the rest 'anudātta' suggests a virtuous performance, not elsewhere.'

The intrinsic validity (svataḥprāmāṇyam) of perpetual Śrutis (Vedic comprehension) is deliberated in saying afterwards.

smṛtayo bahurūpāśca dṛṣṭādṛṣṭaprayojanāḥ/
tamevāśritya liṅgebhyo vedavidbhiḥ prakāśitāḥ// (Vākyapadīya Brahmakāṇḍa, 7).

Multiform law texts; Smṛtis concede apparent but also unforeseen objectives, enforcing the hetus (middle terms)—the Vedic scholars dwell upon their cognizance.

Sūryamāyā: Vedic seers (ṛṣis) adhere to Śruti or Veda (revealing genesis), testimony-in-itself, in contributing a good number of Smṛtis (derivative authority) in light, does considerable argument about "sentence content" for direct and indirect fruition; mundane happiness other than the unobserved. The source of one division holds references to the mainspring Vedic dictum that is available; on the other hand, it lacks rather ritualistic learning in conduct. Smṛtis occur in Āyurveda, Śukranīti, and Kāmasūtra for worldly pleasure—indoctrinating health benefits, social norms, besides the family system, prosperity's acquisition, succeeding children, etc. Audible-aphoristic-composition (sūtra); Gṛhyasūtra, Dharmasūtra propositions: daily performance (nitya karma), desire-fulfilling (naimittika karma), apart from reward-seeking sacrifices (kāmya karma) to purify the inner instrument-of-knowledge (antaḥkaraṇa), i.e., mind (manas) et al., what's more release (apavarga) Additionally, Manusmṛti notices both purposes.

Smṛti undertakings of sages are not their personal imagination instead of authoritative verification (pramāṇa) that retains sound cognition (Veda), unlike the unauthorized (apramāṇa); insipid inferences such as āgama of the heterodox'.

Remarks

The reality of language thus correlates to and mirrors existent beings' intelligible stations in that flux, which is together speakable (perceptual) over and above conceptual.

References

Bhattacharya, Bishnupada. (1985). Bhartṛhari's Vākyapadīya and Linguistic Monism. Bhandarkar Oriental Research Institute, Poona.

Commented by Mishra, Jwalaprasad. (Samvat 1964). Kālidāsa' Raghuvaṃśamahākāvya. Shri Venkteswar Steem Press, Bombay.

Edited by Dharmadhikari, N.T. (1981). Taittirīya Saṁhitā (vol-II, part-I). Vaidika Sanshodhana Mandala, Poona.

Jha, Vedanand. (2002). Comprising commentary Ānandā in Sanskrit and Sūryamāyā in Hindi over Vākyapadīya Brahmakāṇḍa of Bhartṛhari. Mandakini Sanskrit Vidvat Parishad, Delhi.

Lokeswarananda, Swami. (1998). Translated and with notes based on Sankara's commentary: Chāndogya Upaniṣad. The Ramakrishna Mission Institute of Culture, Calcutta.

Shastri, Sudarshan. (1954). Commentary Adhikaraṇasārāvalī on Vedāntasāraḥ of Rāmānuja. Chowkhamba Sanskrit Series, Banaras.

The Collins English Dictionary. (2024). Internet Source.

Translated into English by Sastri Mahadeva, A. (1903). The Taittirīya Upaniṣad. The G.T.A. Printing Works, Mysore.

Translated by Karan Singh. (1987). Muṇḍakopaniṣad. Bharatiya Vidya Bhavan, Bombay.

Jadavpur Journal of Languages and Linguistics
ISSN: 2581-494X

Navigating Linguistic Shifts: An Empirical Exploration of Internet-Driven Semantic Changes in College Students

Varshita Dhanda[1]
Wilson College, Maharashtra

ARTICLE INFO

Article history:
Received 12/09/2025
Accepted 06/10/2025

Keywords:
semantic change,
internet linguistics,
digital communication,
neologisms,
language evolution,
social media

ABSTRACT

Language is a dynamic medium that continually evolves in response to social, cultural, and technological changes. This paper explores how the Internet plays a significant role in semantic language change, focusing on college students aged 18-23 at Shiv Nadar University, India. These netizens actively participate in the evolution of linguistics, especially the ones shaped by online platforms. This study implements qualitative methods, including online chats, Google Forms and personal interviews, to collect data from 300 students. Etymological and critical discourse analysis methods are used to analyse how words have undergone semantic shifts in meaning, context and usage.

The analysis unveils that the Internet expedites linguistic change by introducing neologisms and reshaping language's formal and informal boundaries. Some students consider these linguistic changes to be a part of the growth of language and culture, whereas others have asserted their concern about its impact on traditional language aesthetics and professional communication. The study further underlines the communication gaps within the same generation of certain words that have gone out of circulation or are restricted to specific online contexts or groups.

This study emphasises the complex interdependence between language, culture, and technology, which deepens the understanding of relevant communication trends. It also recognises the room for future development, such as considering the socio-economic peculiarities of online language and the linguistic dynamics within specific online communities. Therefore, by identifying these changes, this study contributed to an improved comprehension of how language shapes and is shaped by the swiftly changing digital world.

Introduction

Language is a dynamic medium constantly evolving in response to social, cultural, and technological shifts. The language reflects a unique socio-cultural identity. With the rise of the internet and social media platforms, a significant transformation has been seen in how people communicate and convey the meanings of words. The younger generation (i.e. Gen-Z), particularly college students aged 18-23, who have grown up in a digital age and are active internet users, are at the forefront of this linguistic evolution. This research paper will investigate how the internet has influenced the semantic language change among college students. The rapid dissemination of information through the internet is causing significant changes in language use.

[1] Corresponding Author Email ID: varshitadhanda@gmail.com

Through an online post or tweet, a new word gets coined or modified and subsequently gets popularised at an astonishing pace, affecting communication and interpretation of language. This linguistic shift must be studied as it reflects critical changes in contemporary norms, trends, and values. The internet is a unique space where language evolves constantly through experimentation and adaptation. On various platforms, new words, phrases, and modes of communication emerge daily, offering insight into the dynamic nature of the human language. Therefore, it is essential to document and analyse these changes thoroughly in the digital environment.

The age group of 18-23 is at the forefront of linguistic change, both as creators and users. Therefore, studying the influence of this change in their language is crucial for gaining valuable insights into the future course of language evolution. This research paper aims to establish that these linguistic changes follow a discernible trend or pattern that requires academic exploration. The significance of this research extends beyond linguistics to educators, communicators, and society as a whole. It demonstrates how language shapes contemporary culture and influences how individuals present themselves in the interconnected world. This study also shows that these semantic changes are not limited to the digital realm but extend to offline communication modes. The research will employ a comprehensive approach, utilising qualitative methods. The paper will explore the semantic changes occurring in language through a specific list of words that have transformed due to the internet and social media. It will investigate how these changes get reflected through shifts in meaning, context, or usage and how they affect college students' communication patterns, both online and offline. To achieve the objectives, categorisation of the sample words will be undertaken in multiple ways. First, words will be categorised based on their method of collection. Two, words will be classified into four categories: words unique to social media platforms and online communication, formal words that evolved in meaning due to the internet, informal words that developed in meaning due to the internet, and words that carry both formal and informal meanings due to semantic changes influenced by the internet.

The research will analyse linguistic innovations, neologisms, shifts in word meanings, and adaptations specific to the internet and social media. The investigation will explore the contributions of various social media platforms, including Twitter, Instagram, Snapchat, WhatsApp and Facebook, to changes in language and semantics. The study will focus on how linguistic shifts impact communication patterns among college students aged 18 to 23 at Shiv Nadar University. The exploration will encompass their interactions, perceptions, and relationships. A random survey of 300 undergraduate students will be conducted to construct the corpus. The data collection will be conducted randomly through various methods, including online chats, personal interviews and Google Forms. This includes evaluating the possibilities and challenges presented by this linguistic transformation. Ultimately, this paper examines how language interacts with the internet, providing insight into how language influences and is influenced by the rapidly changing digital world.

Research Objectives

This paper intends to explore how online platforms like WhatsApp, Twitter, Instagram, and Facebook accelerate linguistic changes through neologisms, abbreviations, and shifts in word meanings. The main objective of this research is to analyze and record how the meanings of words change in the digital landscape and their impact on offline interactions. This research uses semantic, etymological, and critical discourse analysis to scrutinize lexical shifts and the tension between formal and informal usage of words. Additionally, it observes the communication gaps

within the same generation because of varied exposure to internet language. Finally, the study aims to demonstrate how language adapts in an increasingly interconnected world.

Theoretical Background

Language is vital in defining human culture, identity, and society. It is a complex and multidimensional tool that facilitates communication, stores cultural history and represents personal and societal identity. Language is important because it reflects a person's local and global identities. With the advent of the internet, it has become easy to make a local language global and a foreign language local (Menard-Warwick, 2013). The relationship between language, culture, and identity is intricate and ever-changing. Language reflects a person's gender, style, class, and regional and social identity (Labov, 1992). Furthermore, language varies between speakers and listeners, settings, and times (Squires, 2010; Tang, 2022; Labov, 1992). Each factor impacts and is influenced by the other. Language naturally evolves, which further complicates its interplay with other factors. Language is more than just a tool for conveying messages; it mirrors a community's beliefs, values, and social norms. As a part of culture, language has the power to shape it. Language is never culturally neutral, and cross-cultural encounters inevitably highlight the rich points of difference (Menard-Warwick, 2013; Chen et al., 2022). One way this occurs is through the distinct linguistic characteristics of each culture, such as vocabulary, syntax, and even nonverbal communication styles (Max S., 1973). For example, the number of words used to express a particular concept in a language can reveal its significance or complexity within that culture and when compared with another language, there might be and might not be a word for the same concept, reflecting upon the points of differences (Menard-Warwick, 2013; Squires, 2010; Hummon, 1994).

Renowned cognitive scientist and linguist Noam Chomsky significantly contributed to the knowledge of language learning and the underlying neural mechanisms of linguistic proficiency. According to Chomsky's theory of universal grammar, people can learn languages from birth, and the variety of languages arises from the interplay between this universal grammar and particular linguistic input (Chomsky, 2013, p. 651). Identifying universal features in language shows us the shared features that support us in how we think and communicate (p. 659). It helps us see our similarities across various linguistic communities.

Analysing the factors behind language change is equally important. A prominent sociolinguistic figure, William Labov (2002), conducted groundbreaking research on language change and variety. His research focused on understanding the social and cultural reasons behind language change, particularly emphasising sound change in progress. Labov's research on dialectical patterns demonstrated that linguistic features such as pronunciation and grammatical constructs can be linked to socioeconomic factors like ethnicity and class (Labov, 2002, p. 14). Labov also identified two main reasons for language change. One is within the language, which he called "internal factors" (p. 14). This refers to changes in sound patterns or how sentences are structured that slowly alter how words are pronounced. The other reason is social, which he called "external factors" (p. 14). This includes gender, upward social mobility, or borrowing words or expressions from different language communities (p. 8). Internal factors typically occur independently, whereas external factors have a strong interaction (p. 14). Labov also argued that people's speech patterns can be influenced by social factors, even if they are not aware of it. For example, people may start using new words or expressions without understanding the social factors driving the change (p. 9). Other social factors, such as changes in power dynamics or the

emergence of new social groups, can also imp0act language change, adding to the richness and variety of languages (p. 8).

Another prominent linguist, Ferdinand de Saussure (1959), developed a structuralist approach to linguistics, emphasising the intricate interplay between language and culture. Saussure's theory suggests that language derives meaning from the relationships and connections between its signs (Saussure, 1959). These signs comprise the framework of language. He introduced the terms "langue" and "parole" to distinguish between the practical use of language in speech (parole) and its underlying structure (langue) (Saussure, 1959, p. 13). This differentiation is crucial in comprehending how language reflects and influences cultural norms and values. The relationship between language and culture cannot be ignored, and one concept that highlights this connection is linguistic relativity, also known as the Sapir-Whorf hypothesis (Kay and Kempton, 1984). The theory posits that the structure of a language can influence perceptions and understanding of the surrounding world, for example, how individuals see and name different colours (Kay and Kempton, 1984). A language's vocabulary and syntax can reflect its users' cultural perspectives and ideals, dictating how certain concepts are expressed (Kay and Kempton, 1984). Language choices, including dialects and accents, can reveal information about one's upbringing, social status, and personal relationships. Language is crucial in shaping personal and collective identities, influencing expression and associations with national, ethnic, and cultural contexts (Kay and Kempton, 1984).

The introduction of technology, especially the internet, has profoundly impacted how individuals communicate in the modern period, which has irreversibly altered the evolution of language. The internet has developed its unique vocabulary, grammar, and language through the influence of various international languages and cultures. Since the internet society is an extension of natural society, the language used on the internet has become an essential tool for studying the development and complexity of society (Tang, 2022). The internet has enabled people of different languages and cultures to communicate more efficiently, leading to more universal language usage. Additionally, it has created a unique online language. While learning an internet language or a second language, people try to connect it with their mother tongue, creating new words and phrases or giving new meanings to existing ones (Chen et al., 2022). The evolution of language in online communication is characterised by informality, brevity, and the use of visuals. Attributes such as acronyms, emoticons, and abbreviations serve as shortcuts to express complex ideas efficiently. These changes also occur because of rapid texting and weak self-regulation, resulting in erroneous words, phrases, and sentences (Thakur, 2021). Despite the benefits of this evolving language, concerns remain about its impact on language competency and the potential erosion of established linguistic norms (Thakur, 2021; Squires, 2010; Tagliamonte, 2016).

People prefer texting language, known as 'texts', over social media because it saves time and is easy to remember (Thakur, 2021). Social media, instant messaging, and online forums have created a plethora of linguistic norms, including emoticons, acronyms, and hashtags. Squires (2010) mentions two terms, Chatspeak and Netspeak, that refer to the informal and abbreviated language commonly used in online communication. Chatspeak is characterised by higher usage of abbreviations, acronyms, and shortcuts (Squires, 2010; Hilte et al., 2020), whereas Netspeak, also explained by Tagliamonte (2016), is a broader term that encompasses not only abbreviations but also informal expressions and styles of internet language (Squires, 2010). People usually use the internet to coin new phrases to describe current social situations and living conditions and for emotional and social sharing, allowing others to experience the same

emotions and become acquainted with real-life occurrences (Chen et al., 2022; Tang, 2022). For example, ‘MeToo’, ‘Black Lives Matter’, ‘FOMO’, or ‘Epic fail’. Platforms like Twitter and Facebook, with character limits and messaging apps prioritising instant communication, have led to a shift toward shorter, more succinct phrases and more informal language usage (Tagliamonte, 2016). It includes unconventional shortening of words like ‘u’ for ‘you’, ‘w’ for ‘with’, and ‘gud’ for ‘good', alphanumeric forms like ‘2night’ for ‘tonight’ and ‘str8’ for ‘straight’, phonetic abbreviations like ‘bt’ for ‘but’, irregular spellings like ‘havin’ for ‘having’, replacing alphabets with numbers like ‘143’ for ‘I Love You’, lowercasing letters even in the initials of sentences, uppercasing all the alphabets, i.e. Allcaps, to give more emphasis to the word or phrase like ‘OH THANK GOD’, repetition of letters like ‘ Wooooow’, repetition of punctuation marks like ‘Congratulations!!!!!!!’ and missing out vowels like ‘ppl’ for ‘people’ (Thakur, 2021; Squires, 2010; Hilte et al., 2020; Kostadinovska-Stojchevska and Shalevska, 2018; Tagliamonte, 2016). Emoticons, for example, facilitate nonverbal communication by enabling users to convey emotions that may be difficult to express in text-based communication, for example, ‘:D’, ‘:)’, and ‘:'(’ (Hilte et al., 2020; Squires, 2010; Tagliamonte, 2016). Similarly, hashtags are an effective tool for organising and categorising information, making it easier to find and share ideas with others, for example, #crisis, #YOLO, and #Blessed (Hilte et al., 2020). The internet has also created new words and phrases, such as "googling" and "selfie," which are now ubiquitous in modern language. Expressions like "LOL," "BRB," and "ICYMI" are now commonplace in everyday speech (Thakur, 2021; Squires, 2010; Hilte et al., 2020; Kostadinovska-Stojchevska and Shalevska, 2018; Tagliamonte, 2016). Utterance chunking is also a common practice online; people often split sentences into multiple short messages, especially youth (Tagliamonte, 2016). This internet culture allows individuals to create and use language, but misunderstandings arise when words are interpreted differently than intended. As various industries, such as science, technology, and medicine, evolve, new terms and jargon are frequently incorporated into language development, such as phishing, blockchain, telemedicine, etc. These linguistic innovations are a testament to language's ability to adapt to the rapidly changing digital landscape, allowing users to express complex emotions, ideas, or cultural references through a concise and easily accessible shorthand. As people become increasingly connected globally, phrases from multiple languages are frequently used in everyday conversation, creating a rich linguistic mosaic.

The internet allows for multimodal expression, combining text, images, videos, and audio, which has led to the emergence of "memes," which are hilarious cultural artefacts that spread quickly across online communities. Kostadinovska-Stojchevska and Shalevska (2018) mention that memes frequently include linguistic play, cultural allusions, and graphic components, demonstrating how language and visual communication are combined in the digital age. They also point out that memes have their own syntax and structure and frequently use text and simple, culturally relevant visuals to communicate complicated thoughts (Kostadinovska-Stojchevska and Shalevska, 2018). They are introducing new linguistic features and modifications to how concepts are perceived and communicated, aiding in the evolution of language. Thus, new vocabulary has been added to the language, conventional grammar and spelling rules have been altered, linguistic diversity has been promoted, and communication practices have changed due to the internet.

The Internet provides both opportunities and challenges. One challenge is the risk of losing linguistic diversity due to language standardisation from international communication networks. However, it also allows speakers of minority languages to communicate and preserve

their cultural legacy, and non-native English speakers can speak without being judged by English speakers (Thakur, 2021). Digital platforms foster linguistic diversity by facilitating the growth of diverse linguistic groups and allowing people to express their identities via language, dialects, and slang (Menard-Warwick, 2013). Artificial intelligence and machine translation have the potential to help people who speak different languages communicate fluently, but they also raise concerns about the loss of cultural and linguistic nuances. Digital communication practices are more relaxed, resulting in a more liberal attitude toward grammar and punctuation (Thakur, 2021; Squires, 2010; Hilte et al., 2020). People rely less on perfect spelling due to autocorrect technologies and predictive text, which is changing linguistic standards (Tagliamonte, 2016). There is an ongoing debate about the use of Internet language. Some argue it is necessary due to technological determinism, while others criticise it for deteriorating the English language (Squires, 2010). Meanwhile, some say that acronyms and abbreviations are not new to the language and have been used in the past by students in their yearbooks or notes; this makes it difficult to determine whether internet language should be considered an evolution or a devolution of language (Squires, 2010).

Generational change is a significant factor in language change, as younger generations bring novel linguistic innovations that may defy long-standing conventions. This phenomenon is especially noticeable in the quick development of informal language and slang, frequently connected to youth culture. Languages are dynamic and adaptable to the changing requirements of their speakers because of the constant cycle of linguistic change. Numerous factors drive modern language shifts; one of the most significant is the extensive internet use by younger generations. College students, in particular, rely heavily on online platforms for communication and information sharing, with social media, instant messaging, and online forums being their primary modes of contact. The ease and fluidity of Internet speech make it easier for linguistic innovations to spread quickly. For instance, the term ‘tweet’ has evolved to refer to sharing brief messages on Twitter rather than just bird sounds. Similarly, ‘streaming’ now relates to the live broadcasting audio and video information over the internet, rather than just water movement. Beyond technology-related terms, the internet has enabled a semantic evolution, where standard terms have slightly changed meaning as they migrate to the digital environment. For instance, "friend" now encompasses virtual social media connections and real-life interpersonal relationships. This linguistic fluidity highlights the interconnection between language and culture and how language reflects the changing nature of human connection in the digital age.

The genesis of slang can be traced back to people's desire to express their individuality and differentiate themselves from the masses. Eble mentions in her work ‘Slang and the Internet’ (2009) that in the case of the English language, slang first emerged in the 18th century as the underworld jargon. At that time, it was considered of little value and was not utilised in formal or written contexts. Even in the 20th century, minimal scholarly research on slang was done and often dismissed as an accidental addition to the lexicon (Eble, 2009). However, slang serves as a linguistic expression of social affiliation and can be viewed as a regional dialect that offers insight into society. Slang encapsulates the excitement and fun of life and is often used contentless and for recreation. Rap music has popularised them, particularly among young people (Eble, 2009). African-American slang has been disseminated through the commercialisation of rap music via radio, movies, and TV. Some slang is unique to a specific group and serves as a marker of their identity, while others are general and can be used by anyone regardless of affiliation (Eble, 2009). According to Sun and Xu in their work titled ‘Tracing Semantic Variation in Slang’(2022), slang differs from community to community.

Different subcultures or groups use different slang, and the semantic variation of slang sets it apart from conventional language (Labov, 1992). Slang has more semantic variation compared to traditional language. Sun and Xu (2022) explain that the semantic variations are mainly due to communication needs and the desire to form a distinct group identity. Slangs not only vary semantically but also change with time and location. Some fade away, some remain slang, and some become respectable words (Labov, 1992).

A contemporary undergrad life consists of multiple social worlds, and college slang is an essential subculture of this life. Hummon states in his work, 'College Slang Revisited' (1994), that undergraduate students use general and college slang, some of which are slang used in every college, and some are local or regional to the particular institution. College slang, which is more of oral, informal, and highly expressive language, is created and used primarily by students as a part of undergrad life to symbolise identification with other undergraduates (Hummon, 1994). It often becomes offensive to those not part of that subculture.

The quickness and informality of Internet communication make it the perfect environment for slang to grow and spread. Slang evolves rapidly, with terms and idioms gaining and losing popularity in the blink of an eye. The internet serves as a linguistic laboratory where new slang is experimented with, improved upon, and spread, showcasing the fluidity of modern language (Brook and Blamire, 2023). People use slang to describe everyday experiences and create a sense of community, helping them to feel like they belong (Labov, 1992). Online slang and language greatly influence youth identity formation and community building (Eble, 2009; Tang, 2022; Hummon, 1994). Adopting particular slang terminology establishes a common language among peers, strengthening their bond and sense of community. Using the appropriate slang at the right moment is crucial in online chatting, even more so than in face-to-face conversations (Eble, 2009). Online communities based on shared identities, interests, or hobbies develop naturally in the digital era, overcoming geographic barriers. Young people can express themselves authentically in these linguistic groups without being constrained by formal language. The linguistic code separating insiders from outsiders is created through Internet slang and language, strengthening the ties that bind these communities. College students often use slang terms to describe different aspects of their life. They may use specific words to refer to their academic pursuits (such as bookworm, ass-kisser, or chemhead), their extracurricular activities (like crewbies, rambo, rocker, or band fag), and their social life (such as hot babe, cool-head, dopeheads, or smoocher) (Hummon, 1994). These terms allow them to express approval or disapproval, classify others, refer to drugs or alcohol, or describe something dull or enjoyable. As a result, these words become unique and meaningful to the students while remaining unclear to adults (Labov, 1992).

The extensive use of informal language, particularly in written communication, is criticised for potentially undermining grammatical and syntactic norms and compromising linguistic accuracy (Thakur, 2021; Squires, 2010; Hilte et al., 2020; Tagliamonte, 2016). Furthermore, communication obstacles between younger and older generations may result from the rapid proliferation of slang and Internet jargon, which can cause generational gaps in comprehension (Tang, 2022). The use of particular slang frequently carries subtle meanings and references that may elude people outside the immediate peer group, so this gap is not just linguistic but also cultural. Online communication's instantaneous nature and brevity impact word choices, promoting succinct language and the use of Internet slang. Furthermore, students tend to bring these informal communication practices into official writing environments because of the informality that the digital world promotes. This change presents difficulties for educators

and linguists since they must manage the tension between the Internet-shaped linguistic landscape and conventional academic norms. Using colloquial terms, acronyms, and informal language in academic writing denotes a break from traditional conventions. This transition does not necessarily imply a reduction in linguistic skills but rather the incorporation of online communication styles into many linguistic situations. While this fusion increases the expressive capacity of communication, it also raises concerns regarding the compatibility of such practices with established academic traditions.

Research Methodology

Data Collection Methodology

The study was conducted at the esteemed Shiv Nadar Institute of Eminence, located in Greater Noida, India, with the primary objective of exploring the impact of the internet on word usage, meaning, context, and communication patterns. Shiv Nadar University is a distinguished academic institution offering diverse undergraduate, graduate and doctoral programs and fostering a co-educational environment for its students. Shiv Nadar University boasts a student body that is both diverse and representative of various socio-economic and cultural backgrounds, hailing from all corners of India. The research sample comprised undergraduate students aged 18 to 23 from different schools within the university - the School of Humanities and Social Sciences, School of Engineering, School of Management, and School of Sciences. The data collection method was convenience sampling, with a population of approximately 3000 undergraduate students considered for the study. The study involved the random collection of data from 300 undergraduate students of all genders. This sample was selected to represent 10% of the total student population at Shiv Nadar University. The respondents answered Google questionnaires, gave interviews, and shared their digital linguistic data for this study. The collected data consisted of 164 words and some abbreviations and acronyms, which were analysed to discern changes in their semantic usage due to the digital intervention.

Methods

Data was collected randomly from students belonging to various academic domains. The task was accomplished using multiple techniques, such as conducting interviews, creating Google Forms, and observing online chats.

- Data retrieval from online resources – The data collection method is to collect data remotely using a digital platform or messaging application: Snapchat, WhatsApp, Instagram, Facebook and Twitter. This approach allows the researcher to engage with individuals and groups through online chat, facilitating the participation of different individuals at their convenience. Interactions are usually recorded through conversational transcripts, providing a comprehensive record of the data collected. This approach offers advantages such as accessibility, diversity, reduced systematic conflicts, and cost savings, making it a valuable research tool.
- Data collection through online questionnaire distribution – Google Forms are used to collect participant feedback. A detailed questionnaire is formed and circulated among the respondents. Respondents could quickly provide input via web or mobile devices, allowing for flexibility in participation. This data collection method is generally preferred because it maintains anonymity and encourages respondents to give honest and truthful answers, thus increasing the reliability of the data collected. Besides,

online questionnaires are flexible and easy, allowing higher response rates and faster data collection and analysis.

- Data collection through on-site interviews – Data collection through on-site interviews involves conducting direct in-person interviews with individuals to gather information and elicit more profound responses. This method allows researchers to ensure accuracy and personalisation in the data collected, which is beneficial when questionnaire responses are incomplete or lack depth. On-site interviews provide an opportunity to engage with participants more personally, enabling the exploration of nuanced perspectives through follow-up questions and clarifications. This approach facilitates the collection of rich and profound data, enhancing the depth and reliability of research findings. Utilising on-site interviews is essential for obtaining comprehensive insights and ensuring the accuracy and depth of data collected in research endeavours.

Ethical Conduct

Before starting the research study, the researcher diligently informed all participants about its aims and objectives. Respondents were asked for consent. The researcher followed established protocols to protect participants' confidentiality. Respondents were assigned a code number and provided information anonymously. Furthermore, the researcher promised that all records would be deleted after one year, and only verbal data would be preserved for future study.

Data Collection Process

Data retrieval from online resources - The online chats were gathered from participants who voluntarily shared their conversations with the researcher. These chats encompassed a variety of groups, including friend circles, peer groups, club or society groups, and event groups. The groups were observed on diverse social media and messaging platforms: WhatsApp, Instagram, Facebook, Twitter, and Snapchat. The online chats also included conversations between two individuals, which were not necessarily confined to groups. By scrutinising these online conversations, the researcher made meticulous observations about the language used by the students, including the types of words, abbreviations, acronyms, and slang utilised. Based on these observations, a certain number of words/terms were documented through this method.

Data collection through online questionnaire distribution - As part of a research study, a questionnaire was thoughtfully crafted and circulated through Google Forms to gather participant responses. The survey featured a mixed pattern format, beginning with demographic questions such as age and gender before delving into eight queries about internet usage and language evolution. The themes in the questionnaire included the influence of the internet on language, semantic language changes, and overall perceptions of these changes. Various response types, such as yes or no answers, statement answers, and multiple-choice options, were employed to ensure ease of use. The primary purpose of the questionnaire was to encourage participants to reflect on the language they use in both online and offline communications. The study focused on undergraduate students from Shiv Nadar University and used a qualitative research pattern. The goal is not to analyse the number of students using certain words but to explore how

language has evolved and its usage has changed. All the responses collected through Google Forms are available in Appendix B[2].

Data collection through on-site interviews - To gather comprehensive data on the topic, one-on-one interviews were conducted. These interviews aimed to obtain interactive and detailed information and were performed exclusively on the Shiv Nadar University campus. Only students who expressed willingness to participate were approached, and the researcher tried to ensure she did not approach anyone who had already filled out the Google form. The interviews were semi-structured, including a framework consisting of prime questions while simultaneously enabling the exploration of additional lines of inquiry based on the respondents' responses. This approach provided flexibility in the research design, which was deemed necessary to gain a deeper understanding of the phenomenon under investigation. The researcher conducted interviews at a convenient time and location for the students and provided an environment conducive to obtaining accurate information. Throughout the interviews, meticulous notes were taken in an Excel sheet to record precise responses and ensure that no valuable information was lost. The Excel sheet is included in the Project report as Appendix C[3].

Questions asked during the interviews:

- What is your current academic year, and what is your major?
- How many hours per day do you spend actively using the internet (excluding time for school or work)?
- Which online platforms do you use most frequently (e.g., social media, messaging apps, online forums)?
- Do you notice any changes in how you use language (e.g., vocabulary, sentence structure) compared to before you started using the internet regularly?
- Can you provide examples of words or phrases you believe have changed meaning due to internet usage?
- Do you believe that internet language influences language use in other domains of your life (e.g., academic writing, professional communication)?
- Do you embrace or resist new language trends that you encounter online?

Data Organisation

The data collected from various sources was organised into a word list in spreadsheet format. The collected data was thoroughly scrutinised, and the corpus was created by selecting words whose meanings, usage, or context were affected by internet usage. The words have been categorised into three groups based on the collection method: words collected from online chats, words gathered through Google Forms, and words collected through personal interviews. Table 1 includes words that evolved in meaning due to the Internet, words unique to social media and

[2] **Responses collected through Google Forms**: For a detailed view of the responses collected through the Google Form, please refer to the following document:
https://drive.google.com/file/d/1Jxl_JluOs8zSvpo9liyZpXrvY5EpDG6W/view?usp=sharing

[3] **Responses collected through personal Interviews**: For a detailed view of the responses collected through personal interviews, please refer to the following document:
Personal Interview Responses_Thesis.xlsx

online communication, and informal words that gained popularity on the Internet. The data was checked for incomplete responses and duplicate responses. Table 1 shows the categorisation –

Online chats	Google form responses	Personal interviews
Friend, Savage, Bae, Lit, Slay, Tea, Swag, Cloud, Catfish, Block, Follow, Like, Swipe, Streak, Tag, Text, Snack, Thirsty, Woke, Tablet, Viral, Meme, Cringe, Hot, Stream, Influencer, Status, Notification, Filter, Thread, Emoji, Profile, Firehouse, Feed, Handle, Spam, Stories, Bud, Jerk, Fam, Flex, Crazy, Banana, Puff, Simp, Nut, Dead, Gucci, Salty, Mugging, Mood, Grind, Sick, Bruh, Squad, Bougie, Alien, Shit, Hangout, Nubi, Dumbfuck, Troll, Netiquette	Slaps, Down for it, Down, Bet, Cool, GOAT, Lit, Viral, Ratio, Meme, Lo-fi, Chill, Game, Flex, Troll, Catfish, Follow, Influencer, Hashtag, Trending, Dead/I am dead, This is sending me, Drip, Cap, Salty, Like, Friend, Vibe, Delulu, Solulu, Zombying/ghosting, Valid, Slaps, Sus, Stream, Mood, FOMO, Giving, Fire, Slay, Low-hanging, Fruit, Bandwidth, Synergy, Core competency, Circle back, Deep dive, Think outside the box, Retard, Sick, Snake, Cold, Cake, Cancel, Woke, Savage, Burn, Roast, LOL, Crying, GG, Mogging, Braining, Religion, Dharma, Yoga, Meditation, Trip, FR, Bruh, Damn, Fetch, Rizz, Yet, Bussin, Bonk, Simp, Cringe, Comment, Tweet, Baby girl, Shit, Whiff, Ship, Dig, Killer, Low-key, Handle, Soo	Flex, Cool, Brother, Coded, On God, For real/FR, No cap, Boomer, Skibidi toilet, Mid, Goofy-aah, Dude, Scam, WTF, Fannum tax, Slay, Girl boss, Bitch, OMG, Tea spilling, Looks maxing, Try Harding, Bro, Lore, Angst, Broody, Hive-minded, Ate, Yassify, Woke, Ghosting, RN, F, OP, GG, Man, Cuz/becoz/bcz, Game, Sus, Bombastic sad eye, That so Reddit, Cunt, IMO, LMAO, LOL, KMS, SMH, Gaslight, Polly, Gorg, Rizz, Cheat code, Beast mode, Smth/sm, WDYM, Serving/giving, Banging, I'm dead, Sadness, Nth, Pookie, Pull up, Queen/king shit, Drop them, Baby girl, Rocking, Kicks, Rocking clothes, What up, OMW, Sxy, Nvm, RN, Pls, PFA, Bruh, Girly pop, Cannon event, Smol

Table 1: *Data organisation based on the method of collection*

Data Analysis Methodology

Etymology

The study of etymology involves tracing words' originality and historical development, including their meanings and how they have changed over time. It looks into where words come from, returning to their first known forms and how they have developed across different languages and cultures (Mailhammer, 2015, p. 2). According to Mailhammer, etymology traditionally deals with words, which may also refer to other linguistic units. The process involves finding out where something came from (p. 3). One could use it for a single sound or text (p. 3). It looks at how the unit's sound, spelling, and usage changed. In etymology, we usually analyse different aspects of the language, including prefixes, suffixes, and roots, to link words with their meanings in other languages. The study of etymology is concerned with a word's phonological form of the word, inflectional morphology, word formation, semantics, and even syntactic information (p. 5).

Also, there may be social factors as well as language structure that help understand lexical etymology better (p. 6). However, this kind of investigation into etymology requires some typological or preferably already available etymological data to make it plausible (p. 5).
Etymology is important to understand language development; it shows how languages change over time. We can learn about linguistic history and cross-cultural interactions by studying word origins. The definition of words can be clarified through this process. Etymology provides hints about the meaning of a word and its usage. Sometimes, the original sense of a term may differ from its current one; thus, etymology helps explain this diversity in meaning formation (Mailhammer, 2015, p. 3). Therefore, it reveals aspects of culture or society that shaped the language as it carries cultural baggage within or reflects certain historical moments or social behaviours associated with those who use that language.

According to the methodology, both the original and new meanings of the words were recorded. The dataset was prepared by merging survey responses, interview responses, and online chats. The data underwent checks for incomplete responses and duplicate responses. All meanings of words with more than one new meaning were recorded, with response formats standardised, and any typographical errors or discrepancies resolved. Various dictionaries were used to find the original meanings of the words, such as the Oxford English Dictionary, Online Etymology Dictionary, Merriam-Webster, Cambridge English Dictionary, and Collins Dictionary, and the new meanings were collected from the respondents.

Critical Discourse Analysis

Critical Discourse Analysis is a mode or perspective of analysis or interpretation of a subject. It focuses on how language shapes our thoughts, beliefs, and actions, particularly in the context of society and power relations (Van Dijk 2). It is not just about the meaning of a word but involves questions about its usage, context, purpose, and by whom it is used. Critical analysis deals with how language affects people, what ideas it fosters, or the stereotypes it promotes; CDA concentrates more on social and political issues rather than ephemeral fashion trends or beliefs (Van Dijk 3). CDA does not constitute a bound theoretical and conceptual framework but depends on the discourse under consideration for its analysis. Different discourses may adopt different theoretical and conceptual frameworks; however, they may have closely associated frameworks, e.g. critical analysis of news reports will require a different approach from school lesson analysis (Van Dijk 3). Critical discourse analysis links power, dominance, and inequality with language, communication, and discourse (Van Dijk 4). It helps to expose hidden meanings or ideas, challenge unjust systems, and encourage social change.

This research uses critical discourse analysis to show how specific online communities or digital platforms determine the meaning of words or phrases, leading to semantic change. CDA allows one to observe how different groups or entities employ language online to assert dominance or influence and how language transforms through different spaces online (Van Dijk 8). It also explains how new online meanings, expressions, and linguistic practices arise. CDA investigates the societal values that underlie modifications in the online language (Van Dijk). In addition, it discusses issues such as gender stereotypes and their changing attitudes about race, identity, and power structures, especially those associated with technology (Van Dijk 8). Take, for example, the term ‘swipe’, which originally meant a physical movement of swinging and then dragging one’s finger over a touch screen. However, it now has various meanings related to dating apps like Tinder, where users swipe left or right depending on whether they approve of someone. As a result of this move, relationships are being commodified by some people while

others have turned dating into an approach similar to playing games using internet sources. Users quickly assess other people's profile information and images when swiping across them all. Gender has something to do with it, too, along with matters concerning sexuality and identity as well. Men swipe right (approve) more frequently, whereas women tend to be more selective. Therefore, Critical Discourse Analysis provides a nuanced understanding of how technological, social, and cultural factors shape language and meaning.

The method involved identifying and categorising key themes, which were further used to analyse and interpret the data. To achieve this, the data was systematically examined to determine any recurring words or phrases. These words were then grouped into thematic categories based on similarities and semantic connections. These categories include Terminology on Social Media and the Internet, Relationships and Dating, Positive Expressions, Negative Expressions, Influence from Mainstream Culture, Emotional Expressions, Internet Behaviour and Etiquette, Gender and Identity, Youth Culture and Trends. Doing so enabled a more organised representation of the data, allowing for a better understanding of the trends and patterns found within the language. Each category was then analysed in detail, and the context of their usage was studied to establish connections between the words emerging from linguistic trends and socio-cultural references.

Discussions

The first part of this chapter focuses on the semantic etymological analysis of the data. We can see from this analysis how the dictionary definition and connotations of words change over time and help us to understand changes in culture and language clearly. Semiotics provides a comprehensive view of semantic shifts, cultural contexts, linguistic borrowings, and intricacies of online communication. For example, the etymology of the word 'friend' shows how it has evolved from the Old English word 'freond' meaning to love; to its contemporary usage including social media friendship in this digital era thus shifting the societal norms as well as interpersonal connections (Online Etymology Dictionary). When trying to find out the roots of terms such as 'meme', it becomes evident that cultural dynamics are at play that have led to widespread internet phenomena ranging from cultural information transmitted by imitation to funny or viral content posted online (Cambridge Dictionary). On top of this, an exploration into origins shows how internet culture influences and borrows from language formation. For instance, 'Skibidi toilet' came from an animated series of YouTube videos called The Skibidi Toilet, featuring singing toilet creatures who want to take over the world. As per the respondents' interpretation, people now use it to describe something evil. With etymology, one can understand online communication better, as exemplified by "ghosting", where a dating term has been transferred into describing behaviour patterns in digital platforms characterised by the sudden end of communication without any explanation (Cambridge Dictionary).

The word 'woke' was initially used to refer to waking up both literally and figuratively, but now it is used to refer to awareness of social and political issues; some people also use it to refer to the leftist ideology, which seeks to achieve social equality and egalitarianism. 'Gucci' originally referred to the prominent luxury Italian fashion brand. Still, nowadays, people, especially the younger generation, use it to denote something of high quality, style or something they understand to be cool. 'Dumfuck' was initially used to insult someone stupid, but it has developed into various contexts; it is used humorously to express emotions like frustration and affection. The word 'coded' for its original meaning is used for some encrypted messages to ensure secure communication. Still, it is also used for someone who speaks in a particular way or

speaks something particular that only that person can say. Therefore, all the words analysed show meaning, context, or usage developments. A thorough analysis of all the words collected during the thesis is done, and it is attached in the Appendix A[4] for reference.

Another qualitative analysis was done. Table 2 shows the data analysis results, categorised into four groups based on their formality and usage patterns: formal words/phrases, informal words/phrases, words/phrases that carry both formal and informal meanings, and new words/phrases. All these words have experienced semantic change due to the internet. Formal words/ phrases are commonly used in formal settings, such as 'cloud,' 'tablet,' and 'stream.' These are mostly related to technology and business environments. Informal words/phrases are commonly used in casual communication, especially on the Internet. Examples include 'bae,' 'woke,' and 'simp.' These often express emotions, opinions or slang. Words with both formal and informal meanings have evolved to be used in various contexts. New words/phrases are those which emerged because of the internet. The most significant number of words fall in the category of words with both formal and informal meanings, which depict the blurring of formal and informal boundaries. Table 2 clarifies that people, especially the younger generation, are inclined toward using informal rather than formal words, as most words fall into two categories - informal words and words with both formal and informal meanings. The choice of formal versus informal language also reflects the social dynamics and power structures. Formal language is often associated with authority and prestige, whereas informal language is associated with familiarity and emotional communication. Many informal words have been added to the category of words with both formal and informal meanings because one of the meanings, either original or new, is used in formal contexts. The formality of a word depends on how it is perceived or interpreted; for example, the word 'slay' may be construed as killing someone or something and can also be interpreted as a sense of personal fashion or style. Similarly, the word 'snack' can be used in formal settings to refer to a small amount of food, but it can also be used in informal settings when it relates to someone you find attractive. Therefore, the meaning of the word varies depending on the context.

Formal Words/phrases	**Informal Words/phrases**	**Words/phrases with both formal and informal usage**	**New words/phrases**
Cloud, Tablet, Stream	Bae, Lit, Catfish, Woke, Cringe, Fam, Simp, Mood, Bruh, Bougie, Shit, Hangout, Newbie, Dumfuck, Troll	Friend, Savage, Slay, Tea, Swag, Block, Follow, Like, Swipe, Streak, Tag, Text, Snack, Viral, Meme, Hot, Influencer, Status, Notification, Filter, Thread, Profile, Firehose, Feed, Handle, Spam, Story, Bud, Jerk, Flex, Crazy, Banana, Puff, Nut, Dead, Gucci, Salty, Mugging,	Emoji, Netiquette, Skibidi Toilet, Fannum tax, Looks Maxing

[4] **Complete Etymological Analysis:** For a detailed etymological analysis of the data, please refer to the following document: Complete Etymological Analysis .docx

		Thirsty, Grind, Sick, Squad, Alien, On God, Coded	

Table 2: *Data organisation based on the level of formality of words*

Critical Discourse Analysis helped better understand the underlying meanings, patterns, and sociocultural implications implicit in the evolved language. From Table 3, it is evident that the Internet and social media have significantly impacted the semantic language change over various domains, like emotional expressions, social interactions, internet behaviour, as well as gender and identity. The thematic categories that were formed and analysed are:

- Terminology on Social Media and the Internet

This set of terms is used for practices and platforms that enable online communication. The use of these words demonstrates the impact that social media has had on language, as words like "like," "follow", or "tag" have become everyday activities in today's era. Furthermore, phrases like "influencer" and "viral" show how significant social media can be for establishing trends and broadcasting online.

- Relationships and Dating:

This category concerns romantic relationships and dating, emphasising the importance of interpersonal connections within youth culture. These present-day dating behaviours exist in expressions like 'bae,' 'ghosting', and 'thirsty,' which are usually discussed in various internet forums.

- Positive Expressions:

These words, expressions, or phrases indicate endorsement, enthusiasm, or satisfaction. These sayings are sometimes indicative of either prevailing moods or trends. For instance, to describe something as impressive or desirable, one will use terms such as 'lit,' 'fire,' and 'queen/king shit', while other phrases like 'circle back;' and 'deep dive' may suggest a more proactive problem-solving decision-making approach.

- Negative Expressions:

Conversely, there is also a part of speech that is used to show dissatisfaction or negativity. The category also features terms like 'cancel', 'salty', and 'gaslight', commonly used to criticise or denounce any undesired conduct or attitude, more so within online interactions.

- Influence from Popular Culture:

This type of popular culture can be seen here and includes words and phrases derived from popular media, entertainment, and consumerism. Words such as 'bougie,' 'Gucci,' and 'rocking' signify fashion trends and lifestyle, while expressions like 'cannon event' and 'Fannom tax' may refer to particular cultural phenomena.

- Emotional Expressions:

Emotional expressions encompass a variety of feelings conveyed through language. It includes idioms that evoke amusement, exhilaration, or even grief. Some examples are 'thirsty' (sexually needy), 'mood' (used to describe someone who does not care what anyone thinks about them), and 'I am dead', which are often used humorously or emphatically to convey strong emotions or reactions to a situation or content.

- Internet Behavior and Etiquette:

This category deals with language and communication norms in online behaviour. It demonstrates how etiquette and social norms are essential to digital spaces, with words such as "troll," "ratio", and "spam" referring to specific online behaviours or phenomena. Understanding and navigating these norms are essential for effective communication and engagement online.

- Gender and Identity:

This category reflects language concerned with gender identity and personal expression. 'Simp,' 'girl boss,' 'baby girl,' 'low-key' – these words convey attitudinal or stereotypical elements relating to gender roles or expressions. To most, simp indicates a derogatory term used on men who seem too submissive. Girl boss encourages women in leadership positions while baby girls may show affectionate feelings. Low-key implies preferring understatement or a lack of enthusiasm. The terms demonstrate the intricate relationship between gender and power within online environments.

- Youth Culture and Trends:

This category reflects the constantly changing youth culture that influences the evolution of language. It comprises various phrases on fashion, lifestyle, and social trends, reflecting young people's diverse interests and preferences in the digital era. Words like 'slay,' 'Gucci,' 'chill,' and 'looks maxing' emphasise style, leisure, and self-expression.

Serial No.	Thematic Categories	Words/phrases
1	Social Media Terminology	Friend, Block, Follow, Like, Swipe, Tag, Spam, Stories, Profile, Handle, Notification, Thread, Emoji, Hashtag, Trending, Viral, Meme, Filter, Status, Firehouse, Feed, Spam, Influencer
2	Relationships and Dating	Bae, Swipe, Streak, Text, Snack, Hangout, Ship, Ghosting, Baby girl, Hookup, Drop them, thirsty.
3	Positive Expressions	Lit, Savage, Slay, Flex, Chill, Game, Flex, Synergy, Core competency, Circle back, Think outside the box, Fire, Low-hanging fruit, Bandwidth, Serving/giving, Rocking, Rocking clothes, Deep dive, Queen/King shit.
4	Negative Expressions	Savage, Catfish, Block, Dead, Salty, Cancel, Woke, Savage, Burn, Roast, Gaslight, Cancel, Dumbfuck
5	Influence of Popular Culture	Lit, Savage, Bae, Swag, Slay, Flex, Stream, Influencer, Viral, Meme, Trending, Thirsty, Woke, Tea spilling, OMG, Bouige, Gucci, Rocking, canon event, Fannom tax
6	Emotional Expressions	Dead, Thirsty, Woke, Mood, FOMO, Sadness, I'm dead, Crying

7	Internet Behaviour and Etiquette	Ghosting, Block, Follow, Swipe, Spam, Troll, Comment, Tweet, Like, Down, WTF, BRB, SMH, Ratio
8	Gender and Identity	Simp, Bae, Baby girl, Queen/king shit, Girly pop, Girl boss, Low-key
9	Youth Culture and Trends	Slay, Flex, Stream, Influencer, Status, Viral, Meme, Trending, Thirsty, Woke, Gucci, Alien, Bougie, Ate, ate and left no crumps, Simp, Flex, Looks maxing, Bruh, Lore, Broody, Hive-minded, Yassify, Boomer, Skibidi toilet, Mid, Goofy-aah, Sus, That so Reddit, OMG, Tea spilling, Skibidi Toilet.

Table 3: *Thematic analysis*

Conclusion

In conclusion, this thesis has deeply explored how the internet significantly impacts semantic language change, portraying the intricate relationship between language evolution, culture and technology. The primary goal was to understand how we speak and understand language in today's world of the internet. The aim was to understand why and how the words change meaning over time, especially concerning online interactions. The researcher collected data through various methods, including online chats, Google forms, and personal interviews. Different research methods were used, including semantic etymological and critical discourse analysis.

Throughout the process of collecting and analysing the data, several insights were gained. One of the key learnings was the dynamic nature of language in the digital age, where words or phrases experience rapid transformations due to cultural shifts and technological advancements. The study also observed the blurred lines between formal and informal language, noticeable in both online and offline conversations. While interviewing the respondents, they were asked if the language they use on the internet affects their communication in professional domains of life. Most responded positively: people must be cautious while communicating or writing in formal domains to avoid using slang, abbreviations, acronyms or informal words in formal settings. The researcher also observed that many people feel uncomfortable using informal or shortened language but still use it to fit into the societal group they want to join. Various words and phrases are observed, with a few students using them while others are unaware of their existence. This reflects the communication gap within the same generation. This practice can also lead the words out of circulation; that is how many words stop existing after some time. Even if students are aware of the internet words or language, many of them refrain from using this language due to their concerns about how this language change affects their cognitive and communication skills, perceiving it as deteriorating the aesthetics of the traditional language. Conversely, other students advocate accepting change as part of language and a natural process. Though many words or phrases have evolved due to online conversation and have been used primarily on online platforms, many words, phrases, short forms, and features of unconventional language have also found their way into offline conversations.

There is much more to explore in understanding how the internet shapes our speech and writing. One exciting research direction is observing how people's socio-economic backgrounds affect their language use on online platforms and how it eventually affects their offline conversations. For example, how different income levels, cultures, and jobs can affect people's choice of words or phrases. Understanding these socio-economic aspects can help us understand how online language reflects and shapes society. It can also help us observe if there are any barriers to online communication based on someone's economic or cultural background. Moreover, exploring how language develops in specific online communities like gaming communities or some professional networks can provide a better understanding of how online platforms shape linguistic norms and identities. Therefore, continued research in this field can help us understand contemporary communication trends and their societal implications, along with the evolving landscape of digital communication in the future.

References

Brook, M., & Blamire, E. (2023). Language play is language variation: Quantitative evidence and what it implies about language change. *Language*, 99(3), 491-530. https://doi.org/10.1353/lan.2023.a907010.

Chen, F., Li, Q., & Yao, H. (2022). Understanding and studying internet culture by the tool of internet language. In *Proceedings of the 2021 International Conference on Education, Language and Art (ICELA 2021)* (Vol. 637, pp. 1045–1049). Atlantis Press. doi:10.2991/assehr.k.220131.189.

Chomsky, N. (2013). Lecture I: What is language? *The Journal of Philosophy, 110*(12), 645 662. http://www.jstor.org/stable/43820808.

Eble, C. C. (2009). Slang and the internet. In G. R. Gonçalves et al. (Eds.), *New Challenges in Language and Literature* (pp. 81–95). FALE/UFMG.

Hilte, L., Vandekerckhove, R., & Daelemans, W. (2020). Modeling adolescents' online writing Practices: The sociolectometry of non-standard writing on social media. *Zeitschrift für Dialektologie und Linguistik, 87*(2), 173–201. https://www.jstor.org/stable/48604369. Accessed 19 Nov. 2023.

Hummon, D. M. (1994). College slang revisited: Language, culture, and undergraduate life. *The Journal of Higher Education, 65*(1), 75–98. https://doi.org/10.2307/2943878. Accessed 15 Oct. 2023.

Kay, P., & Kempton, W. (1984). What is the Sapir-Whorf hypothesis? *American Anthropologist, 86*(1), 65–79. http://www.jstor.org/stable/679389.

Kostadinovska-Stojchevska, B., & Shalevska, E. (2018). Internet memes and their socio

linguistic features. *European Journal of Literature, Language and Linguistics Studies, 2*(4).

Labov, T. (1992). Social and language boundaries among adolescents. *American Speech, 67*(4), 339–366. https://doi.org/10.2307/455845. Accessed 15 Oct. 2023.

Labov, W. (2002). Driving forces in linguistic change. In *Proceedings of the 2002 International Conference on Korean Linguistics*.

Mailhammer, R. (2015). Etymology. In C. Bowern & B. Evans (Eds.), *The Routledge Handbook of Historical Linguistics* (pp. 423–442). Routledge.

Kirch, M. S. (1973). Language, communication and culture. *The Modern Language Journal, 57*(7), 340–343. https://doi.org/10.2307/324652.

Menard-Warwick, J., Dabach, D. B., & Pérez, M. (2013). Local and global identities in an EFL internet chat exchange. *The Modern Language Journal, 97*(4), 965–980. http://www.jstor.org/stable/43651732.

Saussure, F. de. (1959). Introduction. In C. Bally & A. Sechehaye (Eds.), *Course in General Linguistics* (W. Baskin, Trans.). Philosophical Library.

Squires, L. (2010). Enregistering internet language. *Language in Society, 39*(4), 457–492 http://www.jstor.org/stable/40925792.

Sun, Z., & Xu, Y. (2022). Tracing semantic variation in slang. University of Toronto. www.cs.toronto.edu/~zheweisun/files/emnlp22_slang.pdf.

Tagliamonte, S. A. (2016). So sick or so cool? The language of youth on the internet. *Language in Society, 45*(1), 1–32. http://www.jstor.org/stable/43904632.

Tang, J. (2022). The study of internet language functions from the perspective of language variation and change: Case study of 2020–2021 buzz words. *Transactions on Comparative Education, 4*, 22–29. http://dx.doi.org/10.23977/trance.2022.040103.

Thakur, R. K. (2021). Textese and its impact on the English language. *Journal of NELTA*.

van Dijk, T. A. (2001). Critical discourse analysis. https://discourses.org/wp-content/uploads/2022/07/Teun-A.-van-Dijk-2001-Critical_____discourse-analysis.pdf.

Jadavpur Journal of Languages and Linguistics
ISSN: 2581-494X

Code-Mixing in Khasi-English: Patterns, Functions, and Sociolinguistic Implications

Ishita Chowdhury[1]
Central University of Kerala, Kerala

ARTICLE INFO

Article history:
Received 12/09/2025
Accepted 06/10/2025

Keywords:
Loan words,
sociolinguistics,
Khasi language,
language change,
code-mixing

ABSTRACT

This study investigates the sociolinguistic dynamics of codemixing between Khasi, an Austroasiatic language of Meghalaya, and English, which holds a dominant position in Indian education and media. It explores how and why Khasi-English bilinguals fluidly integrate English lexical and syntactic elements into Khasi, focusing on the sociocultural motivations behind this practice rather than just its structural patterns.

Methodologically, the research employs qualitative fieldwork, including recordings of natural speech, participant observation, and interviews. It documents codemixing across various domains such as education, digital spaces, and daily interaction. The analysis pays special attention to the phonological adaptation of English words into Khasi and establishes a clear distinction between codemixing and borrowing.

The findings indicate that codemixing is a strategic communicative tool, particularly prevalent among urban, younger speakers who associate English with modernity, mobility, and a global identity. For this demographic, blending languages is a creative expression of their bilingual reality, not a sign of linguistic deficiency. However, the study reveals significant generational tension, as older speakers and language purists often view the influx of English as a threat to Khasi's cultural integrity. This conflict highlights an ongoing negotiation between tradition and modernity within the community.

Ultimately, the research contributes to contact linguistics by examining a lesser-studied language pair. It advocates for nuanced language policies that acknowledge the value of linguistic hybridity while also supporting the preservation of indigenous languages like Khasi, offering insights for educators and policymakers in multilingual contexts.

Introduction

In multilingual societies, code-mixing, the blending of elements from two or more languages within a single utterance or interaction, is a dynamic and natural outcome of sustained language contact. It reflects the close interplay between language, identity, and social context, and it emerges wherever speakers regularly negotiate multiple linguistic repertoires. This paper examines Khasi–English code-mixing in Meghalaya, a state in north-eastern India, situating the phenomenon within the broader study of language contact and sociolinguistic change. English, as India's principal global lingua franca and a key language of education, media, and governance,

[1] Corresponding Author Email ID: chowdhuryishita081@gmail.com

interacts intensively with Khasi, an Austroasiatic language of the Mon–Khmer branch that constitutes the most westerly member of this family.

The Khasi people, who call themselves Ki Khun U Hynñiewtrep ("Children of the Seven Huts"), form the largest indigenous ethnic group of Meghalaya and are also found in adjoining areas of Assam and in parts of Bangladesh. Historical records and colonial-era missionary writings indicate that Welsh evangelist Thomas Jones introduced the Roman script for Khasi in the early nineteenth century (around 1841), an innovation that profoundly influenced the language's written tradition and facilitated the growth of Khasi literature and print culture. According to the 2011 Census of India, the Khasi population of Meghalaya constitutes roughly 48 % of the state's inhabitants. Prior to the arrival of Christian missionaries, the Khasis practised an indigenous faith known as Ka Niam Khasi; today, approximately 80–85 % identify as Christian, predominantly within the Presbyterian Church, with smaller Catholic, Anglican and Unitarian communities. A small minority, often through inter-community marriage, follow Islam, while a significant number continue to uphold Ka Niam Khasi or Khasi Unitarianism, founded by Hajom Kissor Singh Lyngdoh Nongbri.

Khasi society is notable for its matrilineal system of descent and inheritance, in which lineage and property pass through the female line. The youngest daughter typically inherits ancestral property, yet the maternal uncles (kñi), led by the eldest, act as custodians of family assets and must consent to major decisions. This arrangement does not exclude men from authority; rather, it establishes a complementary balance in which women sustain the lineage while men exercise significant influence in both household and community affairs. The Khasis are sometimes grouped by geographical location rather than sub-tribe: Khynriam in the central plateau, Bhoi in the northern region (Ri Bhoi), Pnar or Jaintia in the east, and War in the southern mountainous area.

The Khasi language, spoken by approximately 1.6 million people in Meghalaya and smaller communities along the Assam–Meghalaya and India–Bangladesh borders, has been recognised as an official language of Meghalaya since 2005. Linguistically it belongs to the Khasian subgroup of the Khasi–Khmuic branch of the Austroasiatic family and is largely isolating in its morphology, though it exhibits limited derivational processes. In May 2012, UNESCO reclassified Khasi as "vulnerable," acknowledging its relative vitality compared with other minority languages of the region, yet warning of pressures from dominant languages.

Within this sociolinguistic setting, Khasi maintains rich oral traditions, a distinctive syntax and a dynamic lexicon, yet the language ecology is changing rapidly. Urbanisation, expanding education systems, and pervasive media exposure in English have intensified language contact and created fertile ground for code-mixing. Khasi–English bilingual speech displays several characteristic patterns: the insertion of English lexical items into Khasi sentences; the adaptation of English phrases to Khasi phonology and syntax; and, at times, the blending of grammatical structures. These patterns vary with speakers' level of bilingualism, the topic and purpose of conversation, and the social domain in which interaction occurs. For example, speakers may prefer Khasi when discussing traditional or family matters, but employ English terms for modern or technical concepts lacking established Khasi equivalents.

The motivations behind Khasi–English code-mixing are both practical and social. Practically, lexical gaps in Khasi often prompt the borrowing of English terms for concepts introduced through modern education, technology, or administration. Socially, code-mixing serves as a marker of prestige, modernity and upward mobility, while continued use of Khasi indexes cultural identity and community belonging. This duality—English as a gateway to

socioeconomic advancement and Khasi as a core emblem of ethnic identity—requires speakers to negotiate their multilingual identities continually.

By examining Khasi–English code-mixing in this broader cultural and historical context, this study contributes to the understanding of language contact, identity formation, and linguistic change in South Asia. It highlights both the challenges and the opportunities of preserving linguistic diversity in an era of globalisation and underscores the need for deliberate efforts to sustain the Khasi language and its cultural heritage while accommodating the evolving realities of a multilingual society.

Code and Code-Mixing in Linguistics

In linguistics, the term code serves as a neutral label for any system of communication—whether a language, dialect, register, accent, or style—chosen by speakers according to context, topic, or purpose. It allows people to interact without the need to classify the medium as a distinct language or variety and highlights the speaker's ability to select the most appropriate linguistic form for a given situation. In multilingual settings, speakers may not only choose a single code but also move fluidly between codes, sometimes within a single conversation or even a single utterance. This fluidity gives rise to code-mixing, a phenomenon in which elements of two or more languages—lexical items, morphological forms, or grammatical features—are combined in the same stretch of discourse without a change in topic. Unlike simple borrowing, code-mixing is typically spontaneous and reflects the speaker's bilingual or multilingual competence as well as their social and cultural affiliations. It often occurs naturally in bilingual communities where languages coexist and interact, enabling speakers to convey meaning, express identity, and negotiate social relationships. Far from being a sign of linguistic instability, code-mixing demonstrates the creative and adaptive capacity of multilingual speakers to draw on the full range of their linguistic repertoire in everyday communication.

Objectives of the Study

This research investigates Khasi–English code-mixing as a case of language contact in a minority language community. It focuses on the structural and sociolinguistic mechanisms through which Khasi speakers integrate English lexical and grammatical material into their speech and on the broader implications for language maintenance and change. The specific objectives are:

- Document and classify instances of Khasi–English code-mixing in natural speech.
- Analyse phonological adaptation, tracing how English words are modified to fit Khasi sound patterns.
- Identify sociolinguistic determinants—such as topic, setting and interlocutor—that condition the use of code-mixing.
- Map the principal domains of occurrence (education, media, everyday interaction) in which Khasi–English code-mixing is most frequent.
- Differentiate code-mixing from borrowing, clarifying how temporary in-conversation mixing differs from long-term lexical incorporation.
- Examine demographic variables, including age, gender and level of education, to see how these influence the type and frequency of code-mixing.
- Assess speaker attitudes toward the use of English within Khasi discourse and the social meanings attached to this practice.

- Contribute to contact-linguistic theory by providing empirical evidence from a lesser-documented language pair.

Purpose of the Study

This study seeks to present a comprehensive, empirically grounded account of Khasi–English code-mixing within the multilingual context of Northeast India, with particular reference to the state of Meghalaya. Khasi, a Mon–Khmer language of the Austroasiatic family and the principal indigenous language of the region, coexists with English, which functions as a language of education, administration, digital communication and upward socioeconomic mobility. The sustained contact between these two languages provides a fertile ground for examining how speakers strategically combine and alternate linguistic resources to express complex social identities.

At its core, the research aims to document and analyse the structural and sociolinguistic dimensions of Khasi–English code-mixing. First, it investigates the phonological and morphosyntactic integration of English lexical and structural elements into Khasi discourse, identifying the adaptations that enable borrowed or mixed forms to conform to Khasi sound patterns and grammatical norms. By observing naturally occurring speech across varied domains—such as classrooms, media interactions and informal conversation—the study evaluates whether these adaptations follow systematic, rule-governed patterns or display context-dependent variation.

Secondly, the study explores the motivations and functions of code-mixing, distinguishing it from long-term lexical borrowing. Whereas borrowing entails the permanent incorporation of foreign words into the lexicon, code-mixing is a more fluid, spontaneous alternation of languages within a single utterance or conversational turn. Clarifying this distinction contributes to a more precise understanding of contact linguistics and bilingual speech production. Particular attention is paid to the sociolinguistic forces—such as the prestige attached to English, the need to express modern or technical concepts, or the desire to index group solidarity—that shape when and how speakers mix languages.

A further objective is to analyse the impact of demographic variables—including age, gender, educational background and occupation—on the frequency and character of code-mixing. Such factors are expected to illuminate patterns of ongoing language change and intergenerational variation in Khasi-speaking communities. By situating these findings within the broader context of multilingualism in India, the study contributes to the understanding of how minority languages negotiate globalising pressures while maintaining their cultural distinctiveness.

Finally, this research addresses a notable scholarly gap. While code-mixing has been examined extensively in major Indian languages such as Hindi, Bengali and Tamil, the languages of Northeast India remain comparatively under-represented in the literature. Khasi, despite its official status in Meghalaya and its approximately 1.6 million speakers, has received limited systematic attention in contact-linguistic studies. By providing detailed empirical evidence and critical analysis, the study advances theoretical debates on bilingualism and contact phenomena and offers insights of practical value to language policymakers, educators and community leaders concerned with the preservation and promotion of Khasi. In doing so, it underscores the dual imperative of sustaining linguistic diversity and cultural heritage while engaging with the socio-economic realities of the twenty-first century.

Scope of the Study

This research delineates the linguistic and sociolinguistic dimensions of Khasi–English code-mixing, a contact phenomenon that has received little systematic attention in previous scholarship. While much of the Indian bilingualism literature focuses on Indo-Aryan or Dravidian languages such as Hindi, Bengali or Tamil in contact with English, the present study examines an Austroasiatic language in sustained interaction with a global lingua franca, thereby extending the geographical and typological range of contact-linguistic research.

The study adopts a broad contextual range for data collection, drawing not only on interviews or formal speech but also on spontaneous everyday conversations, classroom discourse, local media broadcasts and digital communication (including social-media exchanges). This multi-domain approach captures both the patterned and the highly creative forms of code-mixing that characterise contemporary Khasi-English bilingual practice.

A key theoretical aim is to clarify the distinction between code-mixing and borrowing, two processes often conflated in earlier studies. Using criteria such as frequency of occurrence, degree of structural integration and speakers' metalinguistic awareness, the research identifies whether an English element represents a transient code-mixed insertion or a permanently borrowed item within Khasi. This distinction is crucial for understanding how contact phenomena contribute to long-term lexical change.

Beyond structural description, the study investigates the sociolinguistic motivations and identity work associated with code-mixing. It examines how speakers deploy English within Khasi discourse to index modernity, educational capital or social prestige, and how such practices vary across age, gender, education level and occupational group. Code-mixing is treated not merely as a linguistic strategy but as a resource for constructing and negotiating social identities in a rapidly globalising environment.

Finally, the scope encompasses the expanding role of English in digital and educational spaces, where bilingual switching is particularly dynamic. In social media interactions and classroom settings, English is frequently employed to introduce technical vocabulary, maintain engagement or signal cosmopolitan orientation. By integrating these emerging communicative domains, the study provides a nuanced account of how global linguistic forces interact with a minority language and how Khasi speakers creatively adapt to evolving communicative demands while sustaining their cultural and linguistic heritage.

Research Questions

What are the common patterns and types of code mixing found in Khasi English Bilingual speech?

In which social domains (e.g., education, media, daily conversation) does code-mixing between Khasi and English most frequently occur?

What sociolinguistic factors (age, gender, education, urban/rural background) influence the use and frequency of code-mixing among Khasi-English bilinguals?

How do speakers differentiate between code mixing and borrowing in everyday communication, if at all?

How is code-mixing perceived by different generations of Khasi speakers—do they see it as language enrichment, language shift, or a threat to linguistic purity?

How does the use of code-mixing reflect broader trends of language contact, shift, and change in the Khasi-speaking community

Theories and Sociolinguistic Aspects of Khasi–English Code-Mixing

The study of code-mixing in bilingual contexts has been shaped by several influential theoretical models. Myers-Scotton's Matrix Language Frame (MLF) Model (Callahan, 2002) posits that in any code-mixed utterance, one language functions as the matrix language, providing the morphosyntactic frame, while the embedded language contributes lexical insertions. This framework enables researchers to delineate grammatical constraints governing code-mixing and to identify structural patterns across contexts. Complementing this, Poplack's typology (2009) categorizes code-mixing into insertion, alternation, and congruent lexicalization, providing a method to classify the syntactic positions and fluidity of language mixing. Additionally, Gumperz's interactional sociolinguistics emphasizes the interpretive processes of speakers and listeners, suggesting that code-mixing functions as a contextualization cue that conveys social meaning beyond structural convergence .

Linguistic approaches offer complementary insights. Kachru's model of bilingual competence argues that code-mixing arises from speakers' ability to strategically employ mixed codes for communicative efficiency. Khasi-English bilinguals frequently insert English lexical items into Khasi sentences to clarify complex concepts or enhance stylistic nuance. The Lexical Gap Theory similarly explains code-mixing as a response to the absence of appropriate terms in the native language. In Khasi, technological terms such as charger or battery are often borrowed from English. The interference mechanism further explains structural blending, as when English phonemes are adapted to Khasi phonology, reflecting cross-linguistic influence in both lexicon and syntax.

Sociolinguistic factors are equally critical. Code-mixing in Khasi functions as a marker of identity and group affiliation, particularly among younger speakers, who use it to signal modernity, education, or peer solidarity. Urban-rural differences are pronounced: urban speakers demonstrate higher frequency and structural complexity in mixing, correlating with greater exposure to English through schooling, administration, and media. In rural areas, English is more likely to appear in formal or specialized contexts, reflecting limited daily exposure. Such patterns highlight the interaction between linguistic competence and social environment in shaping bilingual behavior.

The motivational dimensions of code-mixing are also significant. Mixing can serve pragmatic functions—such as filling lexical gaps, quoting accurately, or expressing technical ideas—but it is also a social tool, indexing prestige, cosmopolitan orientation, or alignment with global cultural norms. In Khasi-speaking communities, English is often associated with upward mobility and educational attainment, while Khasi maintains symbolic value as a marker of cultural identity and community belonging. This duality fosters a dynamic linguistic ecosystem in which speakers negotiate identity, social affiliation, and functional needs simultaneously.

Literature Review

The literature on Khasi-English code-mixing (CM) and code-switching (CS) in Meghalaya provides a theoretical framework for understanding these phenomena. Scholars like Weinreich (2011), Gumperz (1977),and Poplack (2004) have extensively studied CM and CS with some differentiating between the two while others use "code-switching" as an umbrella term. Milroy and Muysken (1997) emphasize the complexity of terminology, highlighting distinctions such as "intrasentential" (within sentences) and "intersentential" (between sentences), which are widely accepted by researchers like Redouane.

Khasi-English CM often adheres to linguistic constraints identified by Poplack. These include rules governing conjunctions, complementizers, relative clauses, and free morphemes. For example, English conjunctions like "and" can conjoin Khasi noun phrases, while complementizers like "that" are less commonly inserted into Khasi sentences.
Sociolinguistic factors play a crucial role in CM and CS. Increased exposure to English through technology, education, and its social prestige has led to frequent switches in everyday communication. English medium schools, even in rural areas, have heightened the integration of English into Khasi speakers' lives. Functional motivations—filling lexical gaps, quoting accurately, or discussing specialized topics—also drive these practices.

Sound symbolism and expression in Khasi have been a significant yet underexplored area in Austroasiatic linguistics. Nagaraja (2002) provides a foundational study on this topic, highlighting that Khasi, like other Mon-Khmer languages, possesses a rich system of expression used primarily as adverbs to modify verbs and adjectives. These expressions are often ideophones, employed to intensify or nuance meanings, and are highly productive, with speakers capable of creating new forms spontaneously. Notably, Khasi's expressive frequently involves reduplication and vowel alternation to mark distinctions in size, intensity, or emotion—e.g., [u] typically signals ‘bigness’ while [o] and [a] indicate smaller forms. Expressives also deviate from standard phonotactics, introducing unique consonant clusters. Compared to Semai or Lao, Khasi showcases similar features like sound symbolism through vowel alternation but lacks the infixation found in Semai (1976). As Rabel (1963) affirms, mastery over such expressions reflects linguistic creativity and cultural nuance in Khasi communication.

Colonialism and missionary interventions significantly shaped language recognition and identity in the Khasi and Jaintia Hills (2020). The British introduced standardized education and print culture, with missionary Thomas Jones standardizing the Roman script for Khasi, impacting linguistic practices. Census undertakings during colonial rule linked language with community identity, categorizing languages alongside religion and caste. Early Khasi writers used vernacular print to preserve cultural beliefs, fostering ethnic identity, which later influenced socio-political movements and tribal autonomy. Language became both a shared heritage and a marker of difference, highlighting the interplay between standardization and cultural preservation.

Rynjah and Lyngdoh's (2023) research on noun phrases (NPs) in War-Khasi and War-Jaintia, two Khasi language varieties, explores their syntactic and morphological features, highlighting distinctions between pre- and postmodifiers. Their cross-linguistic analysis reveals shared NP constructions alongside notable differences in functions and attributes. Prior studies by Jyrwa (1988), Buhphang (2013), Lyngdoh (2017), and Rynjah (2022) have examined NP construction in Khasi, using descriptive and theoretical methods to uncover intricate structures.

Nagaraja explores agreement in the Khasi and Munda languages, both belonging to the Austroasiatic language family. Khasi exhibits concordial agreement from the Mon-Khmer branch, where pronominal markers on the subject are repeated in the verb. These markers, such as u (masculine singular), ka (feminine singular), and ki (common plural), also function as noun classifiers with non-human nouns. This agreement is evident in subject-verb constructions and relative clauses but does not extend to objects. Some Khasi dialects, like Lyngngam, lack this feature entirely, while others, such as Jirang and Bhoi, place agreement markers after the verb.

Schnoebelen, as a linguist, critiqued Canonici’s subclass model by highlighting the inconsistencies in how personal and non-personal borrowed nouns are classified. For example, while uneasy (nurse) is universally classified in class 1a, nouns like aikido (mosquito) also

appear, which challenges the logical boundary of this categorization. Schnoebelen also identified that phonological misanalysis could influence classification.

Ngcobo introduced a more nuanced corpus-based perspective, suggesting that noun class assignment in isiZulu does not follow a rigid morpho-syntactic rule but is instead. Factors like education, age, and language prestige (e.g., the dominance of English or Afrikaans) shape how and where new words are integrated into isiZulu's noun class system. This is consistent with Hudson's (1986) assertion that grammar and sociolinguistic variables should be examined to reflect actual language use.

Hendrikse and Poulos contributed theoretical insights through a cognitive grammatical framework, arguing that noun classes form semantic networks and continua rather than discrete categories. This is further echoed by Palmer , who sees noun classes as ontological classifiers influenced by worldview and perceptual categorization. These works reveal a growing recognition that loan word classification is a socially and cognitively dynamic process rather than a fixed grammatical phenomenon.

The marginalization of India's Northeast in the national English-language press highlights how journalism, far from being a neutral conduit of information, is shaped by entrenched socio-cultural binaries. Sonwalkar (2010), as a linguist, argues that the New Delhi-based press—often regarded as the national press— constructs a dichotomy of "us" versus "them," wherein the Northeast is perceived as culturally alien and politically peripheral. This exclusion is not solely the result of editorial oversight but is rooted in hegemonic structures determining what constitutes 'newsworthy' content. Journalists operating within these frameworks often unconsciously replicate dominant societal values, leading to selective reporting that sidelines regions outside the mainstream. This phenomenon results in what Gitlin (Todd Gitlin, n.d.) describes as "symbolic annihilation," where entire populations and their issues are rendered invisible in media discourse, further deepening their marginalization from the national consciousness.

Gruessner provides a historical overview of the Khasi language, spoken by approximately 950,000 people in Meghalaya, India. Khasi, belonging to the MonKhmer language family, shares membership with the languages of the Nicobar Islands and is related to the Munda languages within India.

The paper traces the development of Khasi from an unwritten to a written language, detailing early missionary efforts by William Carey and the subsequent adoption of the Latin alphabet by Rev. Thomas Jones. The choice of the Sohra dialect as the standard form facilitated wider understanding across the Khasi Hills. Despite initial challenges, the introduction of the Latin script and subsequent refinements, including the addition of "ñ" and "ï", supported the growth of Khasi literature and education. Early Khasi authors like U Jeebon Roy and U Rabon Singh contributed significantly to preserving cultural identity and promoting the language. Khasi gained recognition in education, becoming a language for High School and later a Major Language at the University of Guwahati . Pandey (2021) introduces Meghalaya's Khasi language and culture, emphasizing the ethnic group's identity as "Children of the Seven Huts."

Simon (1974) rejects the idea that Khasi has a rigid tense system. The thesis also discusses verb particles, auxiliary verbs (including "passive" constructions), and two-verb structures. Verbal extension and derivation through preverbs and affixes are also examined, with a detailed discussion of the attributive preverb ba, which nominalizes clauses and can transform verbal sentences into nominal ones (1974). The study also explores adverbs and their role in enriching the expressive capacity of verbs, highlighting "phonaesthetic" adverbs (1974)

Methodology

This study utilizes a mixed-methods research design to provide a comprehensive and nuanced examination of Khasi-English code-mixing, integrating both qualitative and quantitative approaches. The methodology is structured to capture not only the structural patterns of language mixing but also the underlying motivations, social functions, and identity markers associated with this bilingual practice.

The qualitative component relies on several key instruments. Semi-structured interviews with a minimum of 20 native Khasi speakers, purposely sampled from diverse age, occupational, and educational backgrounds, generate rich data on personal attitudes and contextual usage. This is complemented by participant observation in naturalistic settings like marketplaces and classrooms, which documents spontaneous code-mixing and its conversational triggers, following the principles of Gumperz (1982). Furthermore, a media analysis of public discourse in print, television, and digital platforms reveals how code-mixing is institutionalized and reflects broader societal trends.

To achieve statistical breadth, a quantitative survey is administered to at least 100 respondents. This questionnaire employs Likert-scale items, multiple-choice questions, and demographic variables to measure the frequency of code-mixing, identify the domains where it occurs, and detect correlations with social factors such as age, gender, and urbanity. The sampling strategy ensures representation from both urban and rural settings, allowing for the analysis of regional variation as highlighted in prior work by scholars like Dkhar (2016).

Data analysis is equally integrative. Qualitative data from interviews and observations are subjected to thematic analysis, often aided by software like NVivo, to categorize motivations and social functions. Linguistically, the structural patterns of code-mixed speech are systematically classified using Poplack's (1980) typology—insertion, alternation, and congruent lexicalization—to understand how English elements are syntactically and phonologically integrated into Khasi frames. The quantitative survey data is analyzed to establish distributions and correlations, identifying the social determinants of bilingual behaviour.

This methodological framework is informed by and builds upon previous research in the region, such as Talang-Rao's conversational analysis and the NEHU Linguistics Department's descriptive approaches. By combining these techniques, the study is positioned to explore the intricate interplay between linguistic structure and social practice, revealing code-mixing as a strategic, rule-governed behaviour that is shaped by factors ranging from lexical necessity and pragmatic convenience to the expression of modern identity and urban affiliation.

Data Representation

The research methodology involved collecting naturalistic data through audio recordings of spontaneous conversations, semi-structured interviews, and observations across diverse settings in Shillong and its environs, including educational, professional, marketplace, and social contexts. A purposive sample of participants from varying ages, professions, and educational backgrounds ensured a comprehensive dataset of everyday Khasi-English code-mixing.

The analysis, guided by Muysken’s (2000) typology, categorizes mixing into insertion, alternation, and congruent lexicalization. It also examines the phonological and morphological adaptation of English words into Khasi and quantifies the frequency of these elements across different social domains. Findings reveal that code-mixing is a systematic and socially significant practice, not random linguistic behavior. For instance, the insertion of English nouns and

technical terms is common in educational and professional spheres, often conveying modernity or clarifying complex ideas. Alternation between languages within multi-clausal sentences showcases bilingual fluency, while congruent lexicalization demonstrates the seamless integration of English into Khasi grammatical structures.

A key finding is the linguistic naturalization of English elements. English words are frequently adapted to Khasi phonology, and Khasi morphological rules, such as verb affixes, are applied to English lexemes. This indicates that code-mixing operates within specific structural and sociolinguistic constraints.

Sociolinguistically, code-mixing serves multiple functions. English insertions often signal prestige, education, and urban identity, whereas Khasi maintains cultural and emotional resonance. Conversational triggers for switching include emphasis, clarification, topic shifts, and references to technical or institutional concepts. A clear generational and geographical divide is evident: younger, urban speakers more freely use English to express modernity and global affiliation, while rural speakers exhibit more selective mixing, primarily for lexical supplementation.

In conclusion, the data establishes code-mixing as a dynamic, rule-governed communicative strategy deeply embedded in the Khasi sociocultural landscape. It is influenced by a speaker's proficiency, social background, and the interaction's context, functioning as both a practical tool and a marker of evolving identity. These empirical observations lay the groundwork for interpreting how Khasi-English bilinguals strategically navigate their linguistic resources to meet communicative needs and negotiate social positioning.

Between Couple

Mom: balei phi reach slem
Dad: Hmm, ka traffic mynta ka sngi pat kham busy, had to take a longer route.
Mom: Phi la leit pickup ia u bahbah?
Dad: Ee, I picked him up from tuition. U long tired bha, slept in the car.
Mom: Poor thing. Ka weather ruh mynta ka kham sniew. It's been raining nonstop.
Dad: True. I think we should cancel tomorrow's picnic mo
Mom: Cancel? Haduh! U bahbah la excited bha eh.
Dad: I know, but if it rains like this, bad ka suinbneng ka dum , it's not safe
Mom: Hmm… maybe we can reschedule? Ka Sunday patde ngim free re
Dad: Okay, let's see the forecast first. If ka sngi shai ka long okay, then maybe we can go.
Mom: Oh, by the way, phi la thied ia ka vegetables na market?
Dad: Em , I was planning to go da lah dep dihsha. Want to come along?
Mom: hoid, but tang shu dep thiah u baby. He's cranky today.
Dad: Alright, then I'll go alone. Text me if you remember anything else.
Mom: Okay, and please don't forget ban thied coriander. Last time phi la klet noh!
Dad: Haha, okay okay. This time ngan kynmaw noh.

Between Granddaughter and Grandmother

Granddaughter: Hi, Mei! Nga la wan mo.
Grandmother: Ooo, my sweetie pie! Kumno phi long? Nga la sngew kwah bha ban iohi ia phi.
Granddaughter: Nga biang bha, Mei. And you? Phi la khlaiñ?
Grandmother: Hmm, kham shitom khyndiat. But seeing you, nga sngewbha shisha.
Granddaughter: Mei, nga wanrahda cake, bad u chow. Phi lah bam cake?

Grandmother: Ee, I eat cake. But just thylliang noh, ka sugar ruh ka high.
Granddaughter: Okay, Mei. Nga peit beit. I'll give you just a little piece.
Grandmother: Good girl. Phi la nang kham shem nongrim mynta.
Granddaughter: Mei, ka school test jong nga ka la kut. Phi lah peit lem bad nga?
Grandmother: Ee, bring it here. Let me see. Hmm... aiu kane? Phi don class test shuh?
Granddaughter: Hooid, next week. And I'm nervous, Mei. Subject ka maths ka da heh.
Grandmother: Oi, maths! Back in my day, nga ieit bha ia ka. Come here, ngin pule lang.
Granddaughter: Mei, phi dang kynmaw kumno ban leh problem?
Grandmother: Kumba ngi ong, "Practice makes perfect." Shong noh, ngin sdang.
Granddaughter: Hehe, okay Mei. But first, iathuh story noh shuwa? Nga ieit ia ki jingphawar jong phi.
Grandmother: Hahaha, to ho! Story shuwa, study hadien. Ka tip briew tip Blei ruh ka don bynta.

Between parents and child
Mother: Child, phi la thiah bha ne em in the afternoon?
Child: Hmm... thiah bha, mom, but I had a weird dream.
Father: Kumno ka dream? Ka long scary ne funny?
Child: It was scary! U don u briew ba khmih kum u robot.
Mother: AChild: Mom, I'm not hungry. Nga sngew kdang.
Father: Kam dei ban leit skul khlem bam. You'll feel weak.
Mother: Just eat half, okay? Bam tang ka pylleng bad u ruti.
Child: Okay okay... but can I eat jam instead of pylleng ?
Father: Heh! Lada phi thiah bha, jam ruh lah ban ioh.
Mother: Toh, pynleit jingmut noh. Ka por school ka la jan.
Child: toh, I'm going. Bye mom; bye Papa!
Parents: Bye! Be good in school

Between the shopkeeper and the younger boy
Boy: Brother, phi don chips ne em?
Shopkeeper: Hoid, don hi. Do you want the spicy one or plain?
Boy: Ai da uba spicy. Katno ka long?
Shopkeeper: Ka long ₹10 tangka. Lada wan thied ar packet ngn ai offer, ₹18 tyngka
Boy: Okay, ai ar tylli. Bad ai ruh juice ruh da u orange flavour.
Shopkeeper: Juice long ₹20. So together ₹38.
Boy: Take this pisa, kane ₹40.
Shopkeeper: Haha, here's ₹2 back. Phi la wai skul mynta?
Boy: Hooid, la wai. Ka class mynta ka long boring, sir tang explain da explain.
Shopkeeper: Haha, to pynleit jingmut bha ha class. Later, you'll thank him.
Boy: Yeah yeah, my mom ruh I ong kumjuh. Okay, bye, brother!
Shopkeeper: Bye my boy, wan sa shisien!

Between Two Friends
Friend 1: hi , kumno phi long mynta?
Friend 2: Nga biang . and you? Tired ka skul?

Friend 1: shuh biang hi . English class was too much – full of notes.
Friend 2: Heh, lucky nga la skip class.u Sir ruh u ba strick bha!
Friend 1: True! u just keep talking, no break shuh.
Friend 2: Toh, phi la dep thoh ia ka assignment ne em?
Friend 1: Not yet. Nga pyrkhat ban leh mynta noh, before mom starts shouting.
Friend 2: Same here. But I might chill first, peit Netflix shwa.
Friend 1: Toh, sah og ia nga lada phi kwah help, okay?
Friend 2: Sure, sure. Phi ruh ai phone call lada don problem
Friend 1: Okay bro, nga leit shuwa. See you tomorrow!
Friend 2: Bye! Take care, bro!

Between the student and the teacher
Teacher: Phi wan class da lah dei por mynta?
Student: Sorry sir, nga la sah ha traffic. U bus ruh wan step.
Teacher: Ka por class ka long 9:00 AM. Lada phi wan man ka sngi ha kajuh ka rukom, kumno phi lah ban pynlut por ha jingpule
Student: Hooid sir, nga sngewthuh. Next time, I will try to be early.
Teacher: Kam dei tang ba wan class, but you must also participate. Yesterday ha class, phi shu shong khlem kren eiei.
Student: Nga la try ban understand, but the topic ka long complicated bha
Teacher: If you do not understand, ask. wat sngap jar. That is why I am here for you.
Student: Okay, sir, thank you. Nga kwah ban ask biet shaphang ka passive voice, please.
Teacher: Good. Kumta shong noh, bad nga ai short explanation. Later, I will give you a worksheet.
Student: Thank you sir, nga sngewbha ban ioh help.
Teacher: You are welcome. Work hard, success kan wan.

Analysis and Finding
The data on which this part of the paper is based comes from a short questionnaire, observation and recording periods, and some spontaneous Khasi community speakers.

Code mixing
Maschler's Grin seminar paper(2006) defines code-mixing or a mixed code "as using two languages such that a third new code emerges in which elements from the two languages are incorporated into a structural definable pattern''. Hoffman, in the book Introduction to Bilingualism, identifies a few reasons for people to do code mixing. As Habib (2014:8-9), cites 'these reasons are talking about a particular topic quoting somebody else, being emphatic about something, inserting sentence fillers (interjection), repetition use for clarification, expressing group identity, softening/strengthening request, excluding other people when a comment is intended for only a limited audience, and because of real lexical needs. It defined code-mixing as a reason for people to do code-mixing. People are involved in the conversation and the context where the conversation is taking place.

Bokamba, as cited in Bima (2012), Code mixing defines both concepts thus "code-mixing" is the embedding of various linguistic units such as affixes (bound morphemes), words (unbound morphemes), phrases, and clauses from a cooperative activity where the participants, to infer what is intended, must reconcile what they hear with what they understand.

Type of Code Mixing is divided into two types:

a) Inner code-mixing happens because elements are inserted from an original language with variations.
b) Outer code mixing occurs because element insertion stems from a foreign language. It means that the first language of the speaker, in this case, the national language, is inserted with his or her language (original language) or inserted with a foreign language from the speaker's background.

Types of code-mixing ,they are divided into three main types:

Insertion:

Insertion refers to code-mixing in which elements from one language— such as words, phrases, or morphemes—are inserted into another language's grammatical structure (or morphosyntactic frame) within a single sentence or utterance.

a) /ŋa la leit ˈmaːkɛt mɪnˈnɪn ka sŋi/
I went to the market yesterday.
b) /ŋa kwa ban ˈbaɪ ˈkeɪk ˈmɪn.ta ka sŋi/
I want to buy a cake today.
c) /u la rɪˈsiːv ka ˈlɛt na ka ˈpaɾa dʒɔŋ ŋa/
He received the letter from my sibling.
d) /ka ˈdɔkt la ʧɛk i.a ka haˈpɔh ˈhɔspɪt l/
The doctor checked her inside the hospital.

Alternation

Alternation code mixing refers to switching between two or more languages within a conversation, where the alternation typically occurs between sentences or utterances rather than within a single sentence. It is a form of code-mixing where bilingual speakers alternate languages in their speech to communicate effectively, depending on the context.

Examples:

a)/ŋa la dep tʰɔh i.a ka rɪˈpɔt | bʌt aɪ stɪl niːd tu ˈɛdɪt ɪt/
I have finished writing the report but still need to edit it.
b) /u ˈsam.la u ʃɔŋ ha iːŋ | hi ˈdʌz nt wɑnt tu gɔʊ aʊt t ˈdeɪ/
The boy is staying home; he does not want to go out today.
c) /ka la ˈsdaŋ ʃaɪt | ænd naʊ ɪts ˈgɛtɪŋ wɜːs/
It had already started, and now it is getting worse.
d) /tɔ ŋɪn leit ʃa suˈrɔk | wi kæn baɪ ˈsʌmθɪŋ ɒn ð weɪ/
Come, let us go to the road. We can buy something on the way.

Congruent Lexicalization

Congruent lexicalization refers to code-mixing where two languages share similar or congruent grammatical structures, allowing lexical items (words or phrases) from either language to be inserted freely within the same grammatical framework. This occurs when bilingual speakers use elements from both languages interchangeably within a shared syntactic structure without disrupting grammatical coherence.

Examples:

a)/ tɔ draɪv nɔh ʃa kʰlɑw ban tʃɪl/
Let us drive to the forest to chill.
b)/ŋa ɛnˈdʒɔɪ bʱa ban wɒtʃ ˈnɛtflɪks ha iːŋ/
I enjoy watching Netflix at home.
c)/pʰi la ˈfɪnɪʃ i.a ka wɜːk dʒɔŋ pʰi/
Have you finished your work?
d) /ka la stɑːt nɔh ka klɑs | tɔ rʌn/
The class has already started; let us run!

Lexical Borrowing is only possible in a language contact situation. Therefore, it is crucial to consider the contact situation's sociocultural setting and not only the outcome. Even if the contact is between individual speakers of the languages, in a bilingual situation, the borrowing pattern and the actual linguistic performance reflect the general community's exposure to the other language. The result of the language contact also demonstrates the borrowing language's status and what the process of borrowing implies, not only to individual speakers but also to the language community. Some lexical borrowing, cultural gap-filling borrowing, and just-for-prestige borrowing may naturally be attributed to individual exposure, for instance, through traveling, education, and incoming people from other language communities. However, borrowing in the basic vocabulary is always the result of pervasive cultural contact between speakers of the two languages, regardless of individual social contacts. Heavy lexical borrowing often reflects low prestige of the source language, which can contribute to language shift or loss; conversely, lexical stability indicates cultural significance of concepts.

Loan words
A loanword is a word adopted from one language (the donor language) into another (the recipient language) through borrowing. This means the recipient language takes the word, often with some changes in pronunciation or spelling, and uses it as part of its vocabulary. Loanwords are not translated; instead, they are incorporated either in their original or slightly modified form. This process is common when two cultures interact, especially for new concepts, objects, or practices that do not have Indigenous terms in the recipient language.

IPA	**Meaning**	**Source Language**
/ˈpɛn.sɪl/	pencil	English
/ˈfɔt.ka/	Photograph	English
/teɪ.bɔl/	table	English
/skuːl/	school	English
/dɔk.tɔr/	Doctor	English
/bɪs.kɪt/	Biscuit	English

Khasi loanwords

Loan-Words as a Feature of Language Change in Khasi to English Loan words are a significant feature of language change, particularly in the context of Khasi borrowing from English. As English grew in influence through colonization, education, and modern media, Khasi began to absorb English vocabulary to describe new concepts, technologies, and institutions.

Words related to education ("school," "class"), government ("minister", "office"), and technology ("phone", "computer") are now common in everyday Khasi speech. This borrowing often involves phonological and morphological adaptation to fit Khasi patterns. While some purists worry about the erosion of linguistic identity, others see it as a natural evolution and enrichment of the language. The integration of loan words reflects socio-cultural contact and the adaptability of Khasi in a changing world. Ultimately, loan words bridge communicative gaps and mark the dynamic relationship between languages, showing how Khasi continues to evolve while preserving its core structure and identity.

Khasi Word	**Meaning in English**	**English Contextual Example**
Jadoh	A rice and meat dish	"We had jadoh for lunch at the festival."
Shad Suk Mynsiem	Traditional Khasi dance festival	"The students performed at the Shad Suk Mynsiem last weekend."
Dorbar	Traditional village council	The issue was discussed in the dorbar before reaching a decision."
Rangbah Shnong	Village headman	"The rangbah shnong welcomedmthe officials during the inspection."
Tiew-lak	A specific native flower	The tiew-lak blooms during the spring festival in the hills."

Table 1: *loan-words from khasi into english*

English Words	**Khasi Adaptation**	**Meaning or Usage Area**	**Khasi Contextual Example**
School	Skul	Education	"U khynnah u leit skul man ka sngi." (The child goes to school every day.)
Police	Pulit	Law and Order	"Ka pulit ka la wan tohkit halor kane ka jingjia." (The police came to investigate the incident.)

Table	Teibol	Furniture / Household	"Tob noh ia kata ka kot ha ka teibol." (Keep that book on the table.)
Book	Búk	Education / Reading	"Nga don lai tylli ki búk kiba ngi hap pule." (I have three books to read.)
Doctor	Doktor	Health / Medicine	"U doktor u la ai dawai sha u." (The doctor gave him medicine.)

Table 2: *loan words from english in khasi*

Summary of Findings

This study examines the patterns and sociolinguistic implications of Khasi–English code-mixing and lexical borrowing in Meghalaya, providing both structural and functional insights into how bilingual speakers negotiate their linguistic resources. The analysis demonstrates that code-mixing is pervasive across informal and semi-formal domains—such as family conversations, peer interactions, educational settings, marketplaces, and workplace exchanges—and typically occurs through insertional, alternational, and congruent lexicalization processes, as described in Muysken's typology. English lexical items such as feel, environment, start, and work are frequently embedded within Khasi syntactic frames, revealing the role of English in filling lexical and conceptual gaps, particularly for modern, technical, or abstract concepts for which Khasi lacks direct equivalents. Beyond momentary code-mixing, the data reveal a steady incorporation of English-origin loanwords—such as business, plan, gym, school, and work—which are fully adapted to Khasi phonological and morphological systems and regularly appear with Khasi grammatical markers (e.g., ka school, ka gym). This structural integration indicates a level of borrowing beyond casual mixing and reflects the growing prestige of English, especially among younger urban speakers exposed to formal education, media, and digital communication. Importantly, these patterns suggest that Khasi-English bilingualism is not merely an outcome of lexical necessity but also a vehicle for identity construction, where speakers use English strategically to signal modernity, social mobility, and affiliation with a wider global culture. The findings further highlight that while lexical borrowing may enrich Khasi vocabulary, it also raises concerns about long-term language maintenance: heavy reliance on English lexemes, if unbalanced by active Khasi vocabulary development, can index a shift in linguistic attitudes that historically precedes language attrition. Nevertheless, the resilience of core Khasi lexicon in domains of traditional culture and intimate discourse indicates that code-mixing and borrowing, at present, function more as markers of sociocultural adaptation and bilingual competence than as symptoms of imminent language endangerment. These observations underscore the need for continued sociolinguistic monitoring and for community-driven initiatives to strengthen Khasi vocabulary development in rapidly modernizing communicative contexts.

Discussion and Implications

This study reveals Khasi–English code-mixing as a complex and socially meaningful linguistic practice shaped by historical, cultural, and communicative forces in Meghalaya. The increasing presence of English loanwords in modern semantic domains such as education, health, media, and technology illustrates the community's growing engagement with global networks, while the phonological adaptation of these forms into Khasi grammar demonstrates a high level of

linguistic integration. Code-mixing functions not merely as lexical borrowing but as an active negotiation of identity, prestige, and practicality: younger speakers, particularly in urban areas and digital spaces, use mixed codes to project modernity, cosmopolitan outlook, and educational capital, whereas older generations often express concern over language purity and the possible erosion of cultural values. This generational divide underscores the ideological tensions between cultural preservation and the realities of globalization. The study also clarifies the often-blurred distinction between borrowing and code-mixing—where borrowing represents permanent lexical integration and code-mixing remains context-sensitive and spontaneous—showing that many English terms in Khasi are still in transitional stages. Digital media and educational institutions emerge as key domains where code-mixing is most vibrant, enabling speakers to creatively blend languages for humor, emphasis, or accessibility, while traditional and religious settings remain largely resistant to such mixing. Gender and educational background further influence the frequency of mixing, with females and the highly educated showing greater engagement. Despite these insights, the research is limited by a relatively small, urban-focused sample, the absence of longitudinal data, and the underexploration of children's speech and social media discourse. Nevertheless, the findings challenge the notion of code-mixing as linguistic weakness, instead highlighting it as an adaptive strategy of multilingual speakers who sustain their indigenous linguistic heritage while embracing global influences. This underscores the need for policies and educational initiatives that protect Khasi vitality while acknowledging code-mixing as a legitimate and creative expression of bilingual identity.

Conclusion

This study has examined Khasi–English code-mixing as a dynamic sociolinguistic phenomenon within the multilingual landscape of Northeast India, revealing its significance as both a communicative resource and a marker of sociocultural change. The analysis of naturally occurring speech, supported by interviews and participant observations, shows that the integration of English lexical and grammatical elements into Khasi discourse is neither accidental nor symptomatic of language decay; rather, it reflects the adaptive strategies of a bilingual community negotiating modernity and tradition. Code-mixing operates as a socially meaningful practice shaped by variables such as age, education, and communicative context, enabling speakers to signal prestige, global orientation, and group identity while maintaining ties to their indigenous linguistic heritage. Among younger, urban, and educated speakers in particular, the frequent use of English insertions serves as an emblem of cosmopolitan identity and educational capital, illustrating how bilingualism functions as a tool for social positioning. At the same time, the continued vitality of core Khasi lexicon in domestic, cultural, and ritual domains demonstrates that code-mixing coexists with language maintenance, rather than precipitating immediate language shift. These findings underscore the importance of viewing code-mixing not as a threat but as evidence of linguistic creativity and cultural negotiation. For policymakers and educators, the results highlight the need to support Khasi language development—through curriculum planning, media initiatives, and community-driven projects—so that the language remains robust even as its speakers engage with the global linguistic economy. In doing so, Khasi speakers can preserve the vitality of their mother tongue while embracing the opportunities of a multilingual and interconnected world.

References

Blah, M. L. L. Y. (2020). Language recognition and identity formation in the Khasi and Jaintia Hills. Rupkatha Journal on Interdisciplinary Studies in Humanities, 12(5). https://doi.org/10.21659/rupkatha.v12n5.rioc1s17n2

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology.Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Denzin, N. K. (2017). The Research Act: A Theoretical Introduction toSociological Methods (1st ed.). Routledge. https://doi.org/10.4324/9781315134543

Diffloth, G. (1976). Expressives in Semai: A Study in Sound Symbolism. In Austroasiatic Studies. https://web.stanford.edu/~eckert/PDF/Diffloth.pdf

Dörnyei, Z., & Dewaele, J.-M. (2022). Questionnaires in Second Language Research: Construction, Administration, and Processing (3rd ed.)Routledge. https://doi.org/10.4324/9781003331926

Dr. Esther Talang-Rao. (n.d.). Factors Influencing Codemixing and Code-switching in Khasi and Jaiñtia Hills.https://doi.org/10.33329/rjelal.8.3.291

Eastman, C. M. (1997). Lesley Milroy & Pieter Muysken (eds.), One speaker, two languages: Crossdisciplinary perspectives on code-switching, Cambridge & New York: Cambridge University Press, 1995. Pp. xiii, 365. Hb $59.95, pb $24.95. Language in Society, 26(4), 600–602. https://doi.org/10.1017/S0047404500021096

Gordon D. Dkhar. (n.d.). Code-Mixing in Standard Khasis: A Comparative Study of Urban versus Rural Speech Behaviour in Meghalaya. Indian Journal Of Applied Linguistics.

Gumperz, J. J. (1977). The Sociolinguistic Significance of Conversational Code-Switching. RELC Journal, 8(2), 1–34. https://doi.org/10.1177/003368827700800201

Gumperz, J. J. (1982). Discourse Strategies (1st ed.). Cambridge University Press.https://doi.org/10.1017/CBO9780511611834

Hakmana Pranara Liyanage Waruni Shashikala. (2021).Exploring the Sociolinguistic Aspects of Sinhala-English Code-Mixing among Urban Sri Lankan Bilingual Speakers. Open Journal of Modern Linguistics, Vol.11 No.2.

Halmari, H. (1996). Bilingualism (2nd ed.). Suzanne Romaine. Oxford: Blackwell, 1995. Pp. xvi + 384. $24.95 paper. Studies in Second Language Acquisition, 18(3), 369–370. https://doi.org/10.1017/S0272263100015096

K.S. Nagaraji. (2002). Sound Symbolism and Expressives in Khasi. Bulletin of the Deccan College Post-Graduate and Research Institute, 62/63, 77–91. https://www.jstor.org/stable/42930605

Ngcobo, M. (2013). Loan words classification in isiZulu: The need for a sociolinguistic approach. Language Matters, 44(1), 21–38. https://doi.org/10.1080/10228195.2012.706317

Poplack, S. (1980). Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL: Toward a typology of code-switching. Linguistics, 18(7–8).https://doi.org/10.1515/ling.1980.18.7-8.581

Poplack, S. (2004). Code-Switching. In U. Ammon, N. Dittmar, K. J. Mattheier, & P. Trudgill (Eds.), Sociolinguistics (pp. 589–596). Walter de Gruyter https://doi.org/10.1515/9783110141894.1.4.589

Rabel, L. (1961). Khasi, a Language of Assam. Baton Rouge: Louisiana State University Press

Rabia Redouane. (n.d.). Linguistic Constraints on Codeswitching and Codemixing of Bilingual Moroccan Arabic-French Speakers in Canada. Montclair State University. https://www.lingref.com/isb/4/149ISB4.PDF

Rajendra pandey. (2021). Khasi Language and Culture of Meghalaya. Bahri Publication. https://www.academia.edu/111779029/
Khasi_Language_and_Culture_ of_Meghalaya Rynjah, R. K., & Lyngdoh, S. A. (2023). Cross-Linguistic Comparisons of Noun Phrase Constructions in Khasi Varieties. Indian Journal of Language and Linguistics, 4(2), 42–53.
https://doi.org/10.54392/ijll2325

Shana poplack. (2009). Sometimes ΠΙ starts a sentence in SpanishY TERMING EN ESPANOL: toward a typology of code-switching. DE GRUYTER MOUTON. Simon, I. M. (1974). The verbal piece in Khasi. https://doi.org/10.25501/SOAS.00028690
Teeter, K. V., & Rabel, L. (1963). Khasi, a Language of Assam. Language, 39(2), 341. https://doi.org/10.2307/411221
Weinreich, U. (2011). Languages in Contact: French, German and Romansh in twentieth-century Switzerland (R. I. Kim & W. Labov, Eds.). John benjamins publishing company. https://doi.org/10.1075/z.166

Jadavpur Journal of Languages and Linguistics
ISSN: 2581-494X

The Glorious Spiritual Memory and Musical Space: The Women and the Language Matrix of Bhakti Songs in Challenging the Oppressive Hierarchies

Aditi Bhunia Chakraborty[1]
Heritage Institute of Technology, West Bengal

ARTICLE INFO

Article history:
Received 12/09/2025
Accepted 06/10/2025

Keywords:
language,
culture,
identity,
Bhakti movement,
discourse analysis

ABSTRACT

The relationship between language and identity is one of a very deeply entrenched type that has attracted the attention of scholars across the world since eons. Language performs the job of expression of identity in different fields and it does manifest the true nature of identity across all domains and in the realm of music too, as has been depicted in this work, with regard to Bhakti songs and verses. Since performance is a part of culture, it is undoubtedly a very important task to understand the inbred relationship between language, identity and culture, with special emphasis on Bhakti songs and verses as a part of culture. By attempting a discourse analysis of the women saints of the Bhakti movement, engaging the perception of philosophy of language, the paper has attempted to unearth the underlying cultural postulations and biases associated with gender roles and identities.

Introduction

Beginnings of Bhakti: The Path of Women Saints and Seekers

The Bhakti movement was a key Hindu religious revitalisation movement during medieval India, solicited to bring changes to all sections of society by propagating the practice of piety to attain salvation. Inspired by poet-saints, belonging to some prominent sects like Vaishnavism (Vishnu), Shaivism (Shiva), Shaktism (Shakti goddesses), and Smartism, the Bhakti movement locally developed around different gods and goddesses and was preached through poems and songs using vernacular languages so that the message reached the masses. The poet-saints, championed a wide range of philosophical positions ranging from the theistic dualism of dvaita to pure monism.

Though the nature of the movement was such that many got gradually ensnared into it over a period of time, it can be rightfully claimed that much before the movement emerged world over, the poet-songwriters could create a very strong impression in the minds of the masses. The uncompromising caste system, the intricate customs that founded the convention of worship and the intrinsic necessity to shift to a more gratifying way of worship and salvation possibly impelled this movement. The movement with its ultimate agenda of attaining salvation, with one such agenda being considering everyone equal in the eyes of God, served as a medium for spreading religion and spirituality to the marginalized sections of the society – precisely women, whose religious manifestation was curbed in many ways.

[1] Corresponding Author Email ID: aditi.bhuniachakraborty@heritageit.edu

It was a movement that was not only targeted at individual salvation and a spiritual union with God but also towards socio-religious equality, thereby vesting the Indian women with power, to be 'themselves'. Over epochs, their diverse voices relating to society were recorded through feminist literature; exploring women's relationships with other human beings, nature, with men – other than their family members or husbands, and with the divine.

The women, who remained ostracized through ages and moved much within the pre-decided limits imposed on them, became instrumental in creating a similar course of action along with the poised, conformist behaviour pattern, when they were delegated with little spiritual power. By flouting the rigid patriarchal model, with utter fearlessness, persistence and their devotion to God, they carved an independent space for themselves and declined to be tied down by societal norms. This matri-potestal force, in their own way, did the unspeakable, and established the real strength of a woman's spirit.

It was the path of spirituality that the women embraced to help themselves break away from the societal dictates. The feminine spirituality that emerged out of their alliance with God, came to be deeply shaped by experiences of discrimination and exploitation, something that was bound to happen because the women were easy targets of all kinds of oppression and maltreatment. In an act of flouting such societal dishonour, women saints ranging from Andal in the south to Meerabai in north India, undeniably expressed a denunciation of patri-potestal authority through their compositions and actions.

Research Objective

Understanding the Language Gender Interplay through Lives and Verses of the Women Saints of India:

This work, being an effort to explore principally the gender relation from a linguistic point of view, would specifically look at the compositions of the women sages mentioned above, from the contexts of language use, which probably will depict both sides of man-woman relationship but the perspective will not be a frozen opposition between men and women, Analysis of the discourse pattern, will perhaps help us to get a hang on the nuances of the language, used in their compositions, helping us to have a glimpse of the women who overcame the socially sanctioned, institutionalised oppression in their own way.

Our more specific goal is to understand why gender ascriptions matter, by understanding how they function at the level of linguistic pragmatics. There are two seemingly straightforward answers to the question, why do gender ascriptions matter?

One is that they describe people correctly or incorrectly, and truth matters and the other reason is that people's feelings can be hurt and they can feel disrespected when they are described or referred to differently from how they present themselves.

We believe that while both these are right at times, at times neither of them get at the ethically deep reasons why such linguistic pragmatics matter. Rather, the whole concept of gender ascriptions function to organize social space, and support or undermine people's autonomy, bodily agency, and self-determination within this social space, in important ways. Given to understand the role of gender allocation, we would try to unravel how the women, in discussion, deficient in the male's aggressive bravery, could separate themselves by their serene stubbornness in docile opposition.

Theoretical Background

Probing into the Broad Array of Disciplines of Linguistics, Paving the Root for the Research:

Before digging deep into how these women and their notion of Bhakti drove them into challenging the oppressive hierarchies, we would try to take a brief look into the lives of the four pivotal women mystics of the Bhakti movement, who expressed their radical views on their relationships with God. The social changes they promoted through this were nothing short of revolutionary. We would try to unravel how these women, deficient in the male's belligerent courage, could discriminate against themselves by their tranquil obstinacy in docile resistance. By doing that, we would be able to have a deeper understanding of the lives of the women, who despite harbouring a movement like bhakti, could stridently contest the repressive hierarchies.

The major focus of this work being the 'language' of discourse of the women Bhakti poets, it has taken the effort to throw some light on early endeavours of linguists in analysing discourse. Even though, JR Firth (7-33), the English linguist, urged linguists to study conversation, for there we shall find the key to a better understanding of what language is and how it works, Firth (118-149) asserted that 'the main concern of descriptive linguistics is to make statements of meaning' and he also argued that language is primarily a way of behaving and making others behave and consequently a linguist ought to alarm himself with the 'verbal process in the context of situation'. In typifying discourse structure in relation to functional units, one must meet head-on the problem of grammatical realisation, i.e. how a multiplicity of different functions is realised by the four major clause types, 'declarative', 'interrogative', 'imperative' and 'moodless' (expressions without verbs) and it is consequently difficult for a hearer to correctly interpret which function is intended. A broad array of disciplines of linguistics, to which the roots of research, so far taken into consideration lies, have varying ideas on what constitutes appropriate and acceptable data. As has been seen above, Firth reasoned for a text-based account, Chomsky (94), in contrast, always insisted for a total dependence on intuition. Having talked about some of the early approaches and discourse analysis being a discipline, which can be explored from different perspectives, it can be said that there are some approaches where 'content' is of primary importance as they deal only with the content of the language being used, the themes or issues being discussed in a conversation. Other approaches to discourse analysis pay more attention to the structure of language and also focus on how the meanings in specific contexts get shaped by the structure. Different theories of grammar are used by the different linguistic approaches that are employed for discourse analysis. These diverse methods also hold different views about how to talk about meaning. Feminist critical discourse analysis is one such linguistic approach that we are going to deal with for the present study. The efforts and approaches to discourse analysis so far discussed can well be called 'descriptive' as their goal had been to describe how language works and why they work that way. Cameron, while writing more broadly about feminist critical discourse analysis and also about feminism and linguistic theory, explained that one of her main objectives was to question the whole scholarly objective bias of linguistics and to show how assumptions and practices of linguistics are implicated in patriarchal ideology and oppression (16). There are other attempts and approaches to discourse analysis as well, which are commonly known as 'critical' as in 'critical discourse analysis". Defining the working of language is not the only concern of those approaches but they also intercede in social or political issues, problems and controversies across the world. Beyond their functionality within a specified domain, critical approaches are further concerned with the treatment of social practices both in terms of social relationships and also in terms of their implications for things like status, solidarity and power (power indicates

taken-for-granted gender assumptions and hegemonic power relations discursively produced, sustained, negotiated and challenged in different contexts and communities). One of the biggest concerns of discourse analysis is to see how things are said in different manners to express one's attitude and emotion and how the use of language in combination with other stuff that isn't language helps people to build identities of their own. If one has to establish oneself as a member of a street mob of a certain sort, one has to speak in the 'right' way. One has to engage distinctive ways of thinking, interacting, acting, valuing, feeling, and believing as well. The hawkers in pursuance of their profession, a lawyer in profession, a regular at the local bar and a feminist are found to be exhibiting similar kind of features. Pertaining to the earlier statement, it is appropriate to specify over here that feminists do have their own style of doing and saying things and that calls for hooking up discourse with feminism. The present paper, the major focus of which is to find out the function of the language used by women, (that naturally occurs in their presentations), in depicting gender images, is going to primarily harbour a critical approach of discourse analysis.

Research Methodology

The Poets and Poetries of the Woman Saints of Four Major Regions Serving as the Corpus:

The woman saints, who we would focus on, for the present purpose are Andal, Akkamahadevi, Janabai, Gangasati, and Meerabai. The poets selected for conducting the study, represent the poetry of four major regions of India, written in local languages beginning in the 12th century in South India to the 16th century in North-Western India. An often cited verse maps the birth and spread of the movement.

Bhakti took birth in Dravidian lands,
Ripened in Karnataka,
Came to womanhood in Maharashtra,
And grew crone-like in Gujarat,
Reaching Vrindavan, she re-emerged
A nubile young woman.

(translated by Schelling, Andrew)

(www.Shodhganga.inflibnet.ac.in)

Thus, accordingly, the selection of the poets have been done, ranging from Tamilnadu to Rajasthan, as the untrammelled expressions of the poets produced a rich corpus of literature serving as potential tenets of expression of freedom, facilitating the exploration of their compositions from the contexts of language use.

The panorama of female poet-saints within the Bhakti movement gave rise to a distinct genre of expression in Tamil, Kannada, Marathi, Gujrati, Marwari and so on. While the vachanas of these female saints did reveal the multifarious aspects of Bhakti, the vachana poets placed immense importance on the agreement of speech and action. This harmony, they highlighted, was key to the veneration of Shiva. Thus, vachanas—these intensely passionate dialogues in quest of union with the ishta devata did not only break the existing literary tradition, but also discarded the social divisions, hierarchies, ceremonial structures of worship, and pursuit of worldly gratification. Here, the unity between what one says and how one acts is central to aikya (union) with the Lord. Among these poet-saints, some were extraordinarily radical in their rejection of social norms and values, leaving husbands, families, and society behind, in order to extol their love for God. Others attempted to fit into more traditional roles in society by

maintaining their responsibilities as wives and mothers. All wrote exquisite poetry that has been passed on through bards and singers throughout India.

The Bhakti movement that flourished across various literary cultures, with the vachanas playing a pivotal role, did not develop with the exclusive purpose of turning into a major literary form, the language used by the poet-saints, the content of their vachanas and the people they addressed through these vachanas, broke with the existing literary canon and as a result, brought about a crucial turn in the language and literature of their times. Thus, it is no surprise that this period greatly witnessed the blooming of vachanakaras (vachana poets).

Discussion

Deliberating on the Vachanas (verses) from the Philosophical Aspects of Language:

A thorough scrutiny of the vachanas has revealed over and over again that each poet had their own language of allusion while throwing back on the social and spiritual realities. For instance, Andal's poetries are steeped in mystical veneration, Akkamahadevi's vachanas are bewildering in their lyrical expanse, Janabai's abhangs are expressions of bold imagery, and Meerabai's verses are symbolic of eroticism and sublimity, filled with pathos and longing to see and experience her Govardhana. All the poets, marginalised along the lines of caste, class and gender have come up with such distinct expressions that it is important to focus individually on each of them. The paper, technically being an attempt to analyze the poetical works from a sociolinguistic point of view, would concentrate on the aspect of discourse analysis, analyzing the texts from the perspective of speech act theory. The texts selected for the purpose are translations of the verses composed by the women saints in their own languages. To start with, we would focus on Andal, one of the greatest precursors for freedom of spiritual expression.

There is a very different kind of myth associated with the birth of Andal. It so happened that Vishnuchitta, a devotee of the Lord Vatapatrashahi, while digging around his garden, found a large box. When he opened the box, to his utter astonishment, he found a beautiful glowing girl smiling up at him. Believing that this was ordained, Vishnuchitta brought the baby home, raised the child and named her Goda or Kodai. Goda was no ordinary girl, but someone, who was beloved of the Lord himself. Her spiritual magnificence was such that the Lord himself wished to share her presence. From that day onwards, Vishnuchitta started calling her 'Andal', the girl who ruled over the Lord. Andal's devotion and love for her God grew relentlessly with time. She was merely twelve to thirteen years of age, when she was well versed in the mythologies. As a kid, she used to go around motivating her friends extolling the virtue of the Lord, while undertaking penance to attain the mercy of the Lord. On visiting the temple, the Lord would give the girls his darshan, his beauty would stun them and their verses would pour out with eloquent praise for the splendour of the Lord. After blossoming into a beautiful young woman, Andal declined to marry anyone else but Lord Ranganatha, the reclining form of Lord Vishnu, who resided at the great temple town of Srirangam. She was so deeply in love with the Lord that one day she fell asleep and dreamt about sending messengers carrying her message of love to her beloved king. In her dreams, she fantasized about her wedding, where a thousand dazzling elephants carrying her love came to her doorstep. On waking, she was pleased with her dreams and thanked Krishna for coming to her. The Lord also was extremely pleased and told her to sing more songs which he would continue listening to. Andal felt blissful being requested by the Lord and in vivid details sang what came to be known as the 'Nachiyar Thirumozhi'—one hundred and forty three mystic poems known as "pasurams" of love and longing. The pasurams (verses) are organized in fourteen segments and each segment is called a thirumozhi. Each thirumozhi deals with one

specific topic. Among the several thirumozhis, we will, for the present purpose, take into consideration some verses from the first, fourth and sixth in order to express the radicality of the mindset of a woman of that time. The first is a set of verses called 'Varanam Aiyiram', consisting of ten sub-verses. It relates the sequences of her dream of getting married to Narayana or Vishnu. The verses in translation read:

Tomorrow is the Muhurtam, It has been thus decided
The tender leaves sheathing areca nuts have been weaved ,
to form a canopy over the marriage pandal,
Into this entered a strong youth bearing the names,
Narsimhan, Madhavan and Govindan,
Friend! So I saw in my dream!
(https://kcsujata.blogspot.com/2011/04/varanam-aayiram-by-sri-andal-verses-and.html)

Then again she chants
Indra came along with other Devas and sought my hand for Krishna,
They discussed the details of the marriage ,
Durga ,my sister in law,
Draped me in the bridal saree and decked my hair with scented flowers.
Friend! So I saw in my dream!

(translated by K. C. Sujata)
(https://kcsujata.blogspot.com/2011/04/varanam-aayiram-by-sri-andal-verses-and.html)

In the next part of her verse, in a bid to describe the rituals of her marriage, she says
From all four directions Brahmins brought sacred water,
Sprinkled it all over the place, Chanting the Vedas loudly and clearly,
And bound me with a sacred thread to the Pure One wearing a garland of flowers,
Friend! So I saw in my dream!

(translated by K. C. Sujata)
(https://kcsujata.blogspot.com/2011/04/varanam-aayiram-by-sri-andal-verses-and.html)

The description of the event continues and she chants
Mridhangam was beaten,
Conches having two stripes were blown for a long time,
My would be husband the handsome Madhusudhanan came to the pandal,
That was decorated with strings of pearls,
And took my hands in his!
My dear friend! So I saw in my dream!

Mantras were chanted by priests,
Who recited them clearly with good wishes in their hearts,.
The sacred grass with green leaves were spread and sacred twigs [used in hommam] were laid on them
One who is as proud as a angry elephant clasped my hands tightly and strode majestically around the sacrificial fire with me in tow,
My dear friend ! So I saw in my dream!

(translated by K. C. Sujata)
(https://kcsujata.blogspot.com/2011/04/varanam-aayiram-by-sri-andal-verses-and.html)

After the rituals were performed and the marriage was over, Andal expresses herself in another verse, saying
"Smearing the pastes of kungumam and fragrant and cool chandan all over our bodies ,
Relishing the sprinkling of perfumed water on us, by all those gathered around
I went around the decorated town on an elephant along with him seated by my side!
Friend! So I saw in my dream"!

(translated by K. C. Sujata)
(https://kcsujata.blogspot.com/2011/04/varanam-aayiram-by-sri-andal-verses-and.html)
In another verse of The Thirumozhi, she is also seen to be praying to Kamadeva to get Vishnu as her husband. Andal expresses that she will die if she is married to someone else other than Vishnu. The songs serve as mirrors expressing her burning passion to be united with the Lord. In the poem, Andal has tried to express that her Lord longed for her as much as she did long for him. It is an uninhibited poem of love and wild passion written by a young girl budding into womanhood. Being unabashed of her bodily desires, the poem that she composed has been erotically charged, sensuous and deeply felt. It is remarkable for its complexity and frank descriptions of female desire. One such verse of 'Nachiyar Thirumozh'i, where Andal has expressed her love for Vishnu, written with bold sensuality and startlingly savage longing, hunger and inquiry, has been the following. The verse, invoking Kamadev, in translation reads

Tell him I will survive
Only if he will stay with me,
For one day enter me,
So as to wipe away,
The saffron paste, adorning my breasts!

(Translated by. Vidya Dehejia) (Dehejia, 140)
Andal whilst admiring herself wearing the garland which was meant for Vishnu, she further chanted
My surging breasts long to leap to the touch of His Hand
which hold aloft the flaming discus and the bright conch
Coax the world-measurer to caress my waist
To encircle the twin globes of my breasts

(translated by Venkatesh, Arundhati)
The verses deeply expressed her love, longing and frenzy for the Lord while demanding his caress. Erotic in nature, particularly in the backdrop of the time, when women hardly have been gutsy in expressing their emotions, Andal was severely unrestrained in expressing her emotions. If we look at the expressions from a theoretical point of view, principally when the agenda of the work is attempting a discourse analysis, engaging the perspective of philosophy of language, we will find that scrutiny of direct discourse by a character proffers a systematic framework for spotting the unspoken presuppositions, implications, and its effects. In this context, it is probably important to mention that discourse analysis is a process of identifying and discerning social interactions. It is a discipline, which can be explored from different perspectives. There are some methods where 'content' is of primary significance as they take into consideration only the content of the language being used, the subject matter or issues being talked about in a

conversation. Other approaches to discourse analysis pay more attention to the structure of language and also focus on how the meanings in specific contexts get shaped by the structure. Different theories of grammar are used by the different linguistic approaches that are employed for discourse analysis. These diverse methods also hold different views about how to talk about meaning. Gender specific critical discourse analysis is one such linguistic approach that we are going to deal with for the present study.

Aligning with the concept of feminist critical discourse analysis, this paper will try to analyse the discourse pattern of women by employing the philosophical aspects of language, in particular, the speech act theories proposed by J. L. Austin and John Searle and also by using the notion of conceptual metaphor introduced by George Lakoff in treating meaning. Once mentioned the names of Austin and Searle, it should be noted contextually that J.L. Austin's theory was further developed by an American philosopher John Searle who considered the degree to which utterances are said to perform 'locutionary' acts, 'illocutionary' acts, and/or 'perlocutionary' acts. Searle also was responsible for devising a system of speech act categorization where he suggested five illocutionary points that speakers can achieve on propositions in an utterance, namely: assertive, commissive, directive, declaratory and expressive.

The corpus selected for the present purpose consists of such expressions manifesting the interlocutions, involving the speech acts that assert women supremacy. In this work, such women characters and certain expressions they engaged in, defying the stereotype mold for the creation and enlargement of personal space, have been primarily taken into consideration. Since this is a work based on sociolinguistics and also because sociolinguistics regards language to be communicative, social and interactional by nature, it is going to address the issue of conversational analysis in terms of speech acts and also going to deal with how language is shaped and reshaped in the discourse of the lives of the mystics and how it has been instrumental in reflecting and creating the social realities of life. In terms of structure, the vachanas are distinct from other religious literature and literary forms.

Probing truly into the sayings of Andal, as has already been referred to above, in terms of speech act theory, we will find that most of her utterances, mentioned above, are both 'locutionary' and 'illocutionary' in nature. 'Locutionary act' is the meaning of the words used. 'Illocutionary' act is the speaker's intended function and 'perlocutionary' act is the listener's interpretation of the words.Since the very act of uttering is locutionary, so in that way all sayings are locutionary. Illocutionary act refers to the type of speech act that is being performed and the function the speaker intends to fulfil. Thus, after scanning through her vachanas, it can be construed that most of her vachanas are locutionary and perlocutionary in nature. The verse in which she says "Coax the world-measurer to caress my waist, to encircle the twin globes of my breasts", she employs perlocutionary speech act because if viewed at the level of its psychological consequences, it is implicative of the fact that the Lord should undertake the task of coddling and fondling. The speech acts of Andal can also be analyzed in terms of the classification suggested by John R. Searle, who classification was basically an extension of Austin's explanation of the 'illocutionary' act saying, a speech act is often meant to refer to exactly the same thing as the term 'illocutionary' act. Searle's work on speech acts is understood to further refine Austin's conception. If we analyse the speech acts, going by Searle's classification, then we will find that many of Andal's verses are representatives, assertives, commissives, expressives and declaratives. Though, her verses, with regard to her marriage, she claims have been expressed in her dreams, but many of the verses related to the marriage have

truth value. In the verses portraying her sayings in Varanam Aiyiram, she is committing herself to some future course of action, that is the action of getting married and settling down with the Lord (commissives) and through these verses, she is also expressing her attitudes about objects and facts of the world, thereby giving rise to expressive speech act.

In one of the vachanas "Tell him I will survive……" as mentioned earlier, she has been imagining that her marriage would take place with the Lord and Andal while invoking the Kamadev, unabashedly expresses her bodily desires in a very sensuous manner and the verse she thus comes up with in such a context is assertive and directive as well. Assertive because the verse aligns to represent how her feelings have been in the world that is all engulfed with the Lord and directive because they make an attempt to get hearers to do something. In the same verse "Coax the world-measurer to caress.....", Andal uses directive because it makes an attempt to get the hearers to do something. All the verses composed by Andal are, by their own merit, worthy of proving her supremacy that also seem to ignore respect or adherence to a codified ethical system of values of the country at that time. She has thus justly proven herself in upholding a 'reverse social order', much against the prevalent patriarchal system of that time.

Our next woman mystic for consideration here is Akkamahadevi. Akkamahadevi is a 12th century Kannada poet cum sage who has been famous for her vachanas. Creatively sculpted through her writings, her poems are very strong in feminist content. Though, women power has always been accommodated into patriarchal culture through religion, but women sages like Akka have been barefaced in declaring her love for Lord Shiva as rightful, and considered her husband and other family members, as people who jeopardised her zeal to be in harmonious union with her Lord. In her vachanas, she has articulated her radical account on women's position in a patriarchal society. All her fondness towards the Lord, her impudent approach in renouncing her family and all material desires for being in unison with the Lord are expressed through her vachanas. Some of her prominent compositions are cited below in relatable English translations. As H.S. Shivaprakash puts it, '…Akkamahadevi's vachanas are full of images of a journey through villages, towns and forests. Her keen sensitiveness to sights and sounds of nature make her one of the most magnificent nature-poets in the language… '

Being appreciative of an unfruitful corporeal relationship, she considered Lord Shiva as her husband and she thus ends her verses with her 'ankita' (signature) calling Shiva as Chennamallikarjuna, which A. K. Ramanujan poetically translated as, "Lord as White as Jasmine".

Myths and legends suggest that as a woman, Akka has been a spectacular beauty and her long locks added to her alluring look. A Jain ruler, Kaushika, got ensnared by her beauty and fell in love with her. He subsequently proposed to marry Akka. According to some, the wedding did actually take place but on the conditions that the king would neither touch Akka without her consent nor would he hinder her worship. However, when Kaushika contravened the agreement, Akka left his palace, forsaking all her worldly desires and acknowledged Lord Chennamallikarjuna as her soul mate. While leaving her real husband, she stripped off all her clothing and walked out of the palace naked. The account of the entire event is sequentially represented through her vachanas, which she composed over a period of time, one after another. Right after, she came out nude, and walked towards Kalyana, the centre of Virasaiva saints, where Allama and Basavanna ran a school for kith and kin, Allama did not accept her at once. When Allama asked the deserted woman for her husband's identity, she replied she was married to Channamallikarjuna. When she was questioned regarding her nudity and the reason behind

covering her body with her tresses, she answered fervidly through another vachana, which translates

Till the fruit is ripe inside
the skin will not fall off
I'd a feeling it would hurt you
If l displayed the body's seals of love.
O brother, don't tease me
needlessly. I'm given entire
into the hands of my lord
white as jasmine.

(translation by. A. K. Ramanujan, 112-113)

After having a good deal of trials, she was acknowledged in the company of saints. From that point of time started the second quarter of her journey to her Lord. She moved around wild in search of Him and then when finally she saw him, her zeal to receive his grace and be in unison with Him increased manifold times and thus composed another vachana that ardently manifested her longing to be bestowed with the blessings of the Lord

What will come tomorrow, shall come today itself for us.
What will come today, shall come now itself for us!
Don't say then - now, oh the good Mallikarjuna!
Meaning:
Oh Lord, do not delay bestowing Your grace.
If You have determined to bless, why tomorrow, why not at the moment ?
I'm yearning for Your grace. Please don't put conditions
my Lord !

(translated)

(https://shaivam.org/to-practise/dont-postpone-the-grace/#gsc.tab=0)

Out of her ecstatic love, she, in one of her vachanas expressed her heart wrenching cry for release from her familial bonds and be in harmony with Him
Like a silkworm weaving her house with love from her marrow,
and dying in her body's threads, winding tight, round
and round, I burn desiring what the heart desires.
Cut through. O lord, my heart's greed, and show me your way out,
O lord white as jasmine.

(Ramanujan,116)

Another noteworthy composition of hers is the following, which expresses how she found it impossible to reconcile her love for Siva with her love for home and husband. The poem in translation reads

I have Maya for mother-in-law;
the world for father-in-law;
three brothers-in-law, like tigers;
and the husband's thoughts
are full of laughing women:
no god, this man.

And I cannot cross the sister-in-law.
But I will give this wench the slip
and go cuckold my husband with Hara, my Lord.
My mind is my maid:
by her kindness, I join
my Lord, my utterly beautiful Lord
from the mountain-peaks,
my lord white as jasmine,
and I will make Him
my good husband.

(Ramanujan, 141)

In another composition, she imagines the Lord as his groom and insists on cutting clean her relations with the whole world of men and asserting the legitimacy of her love for the Lord, requesting the world not to deter her from getting married to the Lord.

Don't hold me.
Don't Stop me.
Let go Of my hand, the hem Of my sari…
Guru became kin,
Linga, the bridegroom
And I, the bride
The whole world knows…
Channamallikarjuna is my husband, O brother,
No one else can be my man.

(H. S. Shivprakash, 32-37)

In another vachana, she imagines getting married to the Lord and describes the rituals that were performed in the celebration of her spiritual marriage. The vachana in translation goes:

Floor of emerald, hangings of gold, pillar of diamond,
dais of ruby - on this they performed the wedding.
Our people performed at my wedding.
Binding the thread - kaMkaNa - in hand, showering
holy grains (akshata), to the Husband
chennamallikArjuna - they performed at my wedding!
(https://shaivam.org/to-practise/the-wedding-with-god/#gsc.tab=0)

If we look at the vachanas , seriatim, as per the occurrence of the incidents, one after another, we see how, after being wedded to God, she brazenly displays her relationship with God, using vivacious sexual expressions. One such vachana in translation goes

Climbing over the bed,
You came to my hand
Glinting and shining
Dazzled in that blaze
Body, mind and vision stilled,
Lost in the pouring pleasure
Of mating with Sive

Swinging and joining
In the play of love
Devoted, shameless
I shall lie with you
Channamallikarjuna
(https://shaivam.org/to-practise/what-else-you-want/#gsc.tab=0)

The way in which Akka's intense aloofness regarding all worldly affairs, the profundity of her love for Shiva, the anguish of being separated from him and the joyous ecstasy of uniting with him are articulated in language, is beyond belief.

If we look intensely into the character of Akkamahadevi, we will find that Akka, as a woman, was insolent in nature and though most of the female characters, in her times, had the capacity to submerge their individual identity in the identity of the family, handful such women, did not symbolise the eternal woman figure. A thorough examination of the discourse pattern of Akka in her vachanas revealed that covertly or overtly women did make use of language in different ways in order perhaps to be loyal to their social positions at times and also at times defying any stereotyped caste. Whatever the reasons were, AkkaMahadevi has been a defiant spirit, who rebelled against worldly constrictions and prescribed gender roles. The language of her vachanas has thus proved that they are intense and sensual in nature, and have been a reflection of her bold and radical spirit.

Historically and sociologically analysed facts suggest that though, at that time, it was patriarchy and its structural framework wherein all the traditions were enacting social laws and precepts, if we analyze the language and the overall discourse pattern of her poems, the language use would surely help us construe that there were instances where patriarchy was not operative at all or was greatly diluted. Holding fast to the principles of various earlier linguists, who believed that men and women differ in their linguistic behaviour, with men being powerful speakers and women playing a subordinate role all through their parlances, our accountability as researchers would also lie in providing such examples where women disturbed the solidarity in their argot and proved themselves extremely powerful, not being the least bashful about the social taboos. One such woman has been Akkamahadevi, who in spite of being traditional in her speech style has been brave enough for the creation and enlargement of personal space within the given structure of gender relations, with its singular specificity of subordination of women.

The statements of Akkamahadevi, if analysed appropriately, in terms of speech act theory of Searle, clearly signify that most of her statements are 'expressive', 'representatives or assertives' as well as 'commissives' acts. If they are analysed according to Austin's theory, then they are forms of 'illocutionary[2] and locutionary[3] speech acts. The statement of Akka in the first vachana, in response to the reason for her nudity, is 'expressive' because it expresses the speaker's attitudes and emotions towards the proposition. It is a 'representative' too because it commits the hearer to the truth of the expressed proposition. Akka by uttering such things did not only express her attitude and emotions but also did commit to the truth of the expressed utterance. When she says "If the fruit inside is not ripe, the outer skin will not have its ripe colour; If you see what causes sexual desire, you might feel hurt: Thinking so, I have hidden it," she is engaging in illocutionary speech acts. As mentioned above, it is also an 'insult', a type of

[2] Dr. Aditi Bhunia Chakraborty, Asst. Professor. Dept.of Humanities, Heritage Institute of Technology, Kolkata
[3]

perlocutionary act because it reveals the effect the speaker (Akka Mahadevi) wanted to exercise over the hearer (Allama)

Let us once again take a look at the following extract from another vachana already quoted earlier

"I have Maya for mother-in-law;
the world for father-in-law;
three brothers-in-law, like tigers;
and the husband's thoughts
are full of laughing women:
no god, this man.
And I cannot cross my sister-in-law.
But I will give this wench the slip"

(Ramanujan, 141)

This part of the vachana is an example of 'illocutionary' speech act because it consists of the propositional content of the utterance. Going by Searle's parameters of speech act, it is again an example of an 'expressive' because it expresses Akka's attitude about the facts of the world. The thought that gets reflected in the vachana, exhibits that at least in the context mentioned above, all the received wisdom about male superiority and grandeur and the quotidian feminine existence have been pushed to the background and the quiet aura of feminine strength has been foregrounded. Akka, when she placed herself in the context of her relationship with her in-laws, considered herself a submerged identity, dutiful to her family members. But in other matters, particularly when devotion and traditional dharma, (wherein one is supposed to remain committed and dutiful to one's family), come into conflict, in her inexorable desire to be in unison with the Lord, she was the major deciding authority. in the household. Simon de Beauvoir too is sensitive to women's sly empowerment through apparent conformity. She observes: 'By making herself a slave, she has found the surest means of enchanting him'. (Beauvoir 538). But Akka's utterance in the later part of the same vachana, where she says

"I will give this wench the slip
and go cuckold my husband with Hara, my Lord" --- if scrutinized correctly, again in terms of speech act theory, will clearly denote that by Austin's theory, it is an 'illocutionary' speech act, but at the same time is an example of an 'assertive' statement. A speaker's purpose in uttering an 'assertive' is to commit himself or herself to the firmness of the uttered proposition. It is also a feature of 'assertives' that a speaker represents himself or herself to believe what he or she utters and, as the utterance reveals, Akka has committed herself to the firmness of her utterance.

One thing that is proven out of this is that women these days could not rewrite the rules that structured their lives; they were extremely creative in crafting a space from within the prevailing gender system that gave them meaning, solace, and dignity. It is precisely because of such structuring that women, with much dexterity crafted places for themselves.(checked)

Our next important woman mystic of the Bhakti movement, who articulated her radical views on her relationships with God, is Janabai from Maharashtra. Around 1263 AD, just as the Bhakti movement was gaining ground in Western India, right at that point, a girl was born in a small village called Gangakhed, on the banks of the Godavari river in Maharashtra. Her parents were devotees of Lord Vitthal of Pandharpur. Orphaned as a child, she took shelter in a temple and despite being later raised in a family who were devotees of Lord Vitthal, she had to endure dreadful exploitation and abuse. Despite all odds, she remained in close association with Lord

Vitthal and for her, nothing but the Lord existed and she puts it in one of her vachanas, which in translation reads

I eat God
I drink God
I sleep on God
I buy God
I count God
I deal with God
God is here
God is there
Void is not
Devoid of God
Jani says
God is within
God is without
And moreover
There is God to spare

(Kolatkar 2010: 159)

While her devotion never dithered and she also experienced His presence all the time, she sometimes rebuked Him for vanishing and articulated her resentment in no diffident terms, cursing him in public.

In one of her strange verses, Janabai also courted controversy by unfolding herself as a whore and as a wandering singer. Her audacious imagery in the verse, implies that she was probably ill-treated by the men in the household. Disenchanted by the deceitful ways of the world, she described herself as a free-spirited woman who has also not been averse to selling her body. The translated version of the verse reads

Caste off all shame
And sell yourself
In the marketplace;
Then alone can you hope
To reach the Lord
Cymbals in hand,
A veena upon my shoulder,
I go about;
Who dares to stop me?
The pallav of my saree falls away
Yet will I enter
The crowded market place without a thought,
Jani says, My Lord,
I have become a slut
To reach your home

(Vir Sarang 1990: 83)

According to Austin's concept of speech acts, the way Janabai articulates herself in the abovementioned sayings, her expressions are reflections of locutionary and illocutionary speech

acts. Further analysis of the verse according to Searle's speech act theory, suggests that the verses contain 'declaratives', 'commissives' and 'expressives'. Particularly, when in one of the verses mentioned above, she says "I eat God, I drink God, I sleep on God, I buy God, I count God, she employs a 'declarative' speech act because in a declarative convention, if a speaker S utters a declarative sentence with propositional content P, then S thereby commits to acting as though she believes P. In another verse, where she says "the pallav of my saree falls away. Yet will I enter the crowded market place without a thought", she is actually committing to her future actions. The verse in its totality is an 'expressive' because it expresses her emotional exuberance entailing the overwhelming love for the Lord.

Our next woman Bhakti saint is one from Saurashtra between 12th and 14th century AD. Like the other Bhakti saints, stories of her life too, were orally passed on to us. According to the records available, she was born into a Vaghela Rajput family in the Saurashtra region sometime around the thirteenth century and was married to a person named Kahalsang or Kalubha from Bhavnagar. A soul committed to spiritualism, Gangasati took to the task of introducing Panabai, her daughter-in-law to the path of Bhakti and thus composed one bhajan everyday for fifty one days to teach Panabai the essentials of the path and then she took to Samadhi.

Gangasati was intensely religious even before her marriage and she did not put in much effort to draw her husband into spirituality, particularly when her husband also was a follower of Nijiya cult of the Bhakti movement. Gangasati's life and family are noticeably different from the stories of the other woman saints. She was happily married and she managed to strengthen her familial ties through Bhakti. Different from other women poets, Gangasati's holiness was not geared to any specific God. In her compositions, she has made no mention of Krishna and Shiva, there has not been any manifestation of lust or craving. It was nirguna bhakti to the singular all-pervading consciousness. The poem (in translation) below, truly encapsulates her love and devotion towards God:

Once you are on the path of bhakti
Never get tempted
And never keep an amity with a biased person;
Enjoy the bliss of solitude
And have on your head the grace of the Satguru.
Once you are on the path of bhakti…
Pilgrimages and vows won't make any sense
And never do what the Satguru should only do;
Never indulge in trivial things
When you know the secret of the soul.
Once you are on the path of bhakti…
O Harimay, when you have known the world
Keep away from the deceitful ways;
Break all the chains of greed
And always busy yourself in the name of God
Once you are on the path of bhakti…
Mandap and gatherings won't make any sense
These are the signs of the ignorant
Remember these words of Ganga Sati, o Paanabai
This is the way to embrace Ram, the God.

Once you are on the path of bhakti…..

(translated by Sahdev Luhar)
(https://www.sahityasetu.co.in/issue47/sahdev.php)

The entire verse is an example of an Illocutionary speech act because it expresses what the speaker intends to do. After a few more such assertions, she articulates her commitment to her devotional practice, employing imagery typical of battlefields:

"my head may separate from my body and roll off
but my torso will keep on fighting;
I have sacrificed my body…"

Analysing the statement by virtue of the speech act theory proposed by Austin, it can be interpreted as an example of illocutionary speech act as it refers to the intended function or purpose behind the utterance, beyond the literal meaning of the words. Analysing it by Searle's categorisation, the statement demonstrates itself as an example of a 'commissive' as commissive is a reflection of committing to doing something.

Another significant vachana by Gangasati in translation reads

You're one, He and you,
That says it all.
Sup this secret nectar,
In heady blues, dissolve.
Can't have enough of His dapper looks, true,
What is sin, if one's helpless by heart.
His company is the thrill of your song,
Ecstasy lies beyond letters on your tongue.
All illusions shatter to smithereens,
Once you shed your cast.
Meet your true self,
As it descends from heaven upwards.
Contemplating his form is sheer bliss,
Warps or wefts of mind, nothing is amiss.
Says Sati Ganga, listen O Panabai,
Nirvana is your soul's supreme seat.

(translated by Hemang Ashwinkumar)
(https://guftugu.indianculturalforum.in/category/gujarati/)

The verse speaks of a spiritual awakening and union with the divine.

By Austin's deliberation on language philosophy, the above cited vachana, particularly, when Gangasati says "listen O Panabai, Nirvana is your soul's supreme seat", she makes use of perlocutionary speech act, leaving the listener to decide whether to follow the path of abnegation (persuading, prompting action) or not. Following Searl's understanding of the philosophy of language, one might come to a conjecture that the vachana is an example of expressive speech as it expresses their attitudes about objects and facts of the world and also a directive as it makes an attempt to get hearers to do something.

Gangasati composed fifty two bhajans that show the depth of her spiritual awakening. Her stories and verses have been transmitted orally from generation to generation. In most of her vachanas, she has used day to day images which have made her bhajans quite popular.

Our next woman saint of focus is the great Meerabai of Rajasthan, one of the most devout worshippers of Krishna. Meerabai was a 16th century Rajasthani princess, a liturgical songwriter, poet and a mystic who went against the traditions of her times to voice her divine loyalty towards Krishna. Her influence in bhakti poetry has been deep and the innumerable devotional poems and songs, she composed took into account the stories of Krishna's birth, childhood, and play, along with a story about all the ten avatars (reincarnations) of Vishnu. The life of Meera, although substantiated and validated by her own records, is engulfed with myths, but the story at the centre however has been the same, depicting a young girl's love for Krishna.

When Meera was only three years old, she saw an idol of Krishna that was gifted to her father by a sadhu. At the first sight, Meera got deeply enticed by it and since then, she became gritty to make Him her lifelong friend, lover, husband and companion but generally as a convention, when she was a teenager, Meera's marriage was arranged with Bhoj Raj of Chittorgarh. Meera's mother-in-law tried to enforce on her, the worship of Durga as was the custom in their family, but that did not sit well with Meera. Her new family did not appreciate her conduct and they saw her as disobedient and rebellious. Undaunted of what her family expected of her, Meera committed herself to writing verses. Her verses were sometimes erotic, sometimes sublime and passionate, displaying her fondness for Govardhana. One such verse, in translation is the following

I am dyed in the love of Giridhara
My beloved, My companion.
I am playing hide and seek with him
Robed in this smock of five colours
Disguised as I was. He caught me
And on beholding his beauty
I made myself over to Him
Body and Soul.

(Alston 1980: 60)

The example cited is an instance of both 'locutionary' and 'illocutionary' speech acts. By Searle's theory of speech act, it is 'declaratory' because she is doing things in the world at the moment of the utterance solely by virtue of saying that she does.

Another popular verse of Meera, in translation reads

Sister, I have bought Govind in the market,
You only speak of it secretly,
But I proclaim it openly, beating a drum.
You say it was expensive,
But I say it was cheap,
And I measured it out on the scales.
I would offer my body,
I would offer my life,
For a prize beyond all price.
O Lord of Meera, grant her Thy sight
According to Thy promise in her former birth.

(Alston 41)

In the above example, all the three speech acts- 'locutionary', 'illocutionary' and 'perlocutionary' are implanted. 'Locutionary' and 'illocutionary' speech acts need no further explanation but it is also a 'perlocutionary' act because the speaker by uttering "O Lord of Meera, grant her Thy sight" tries to create an impact on the listener, as she is trying to persuade and the listener might act upon it. This is the consequence or result of the perlocutionary act. The cited example is also an instance of 'expressive' and 'declarative'. 'Expressive' because it expresses her attitude about facts of the world and it is also a 'declarative' speech act because the speaker utters the declarative statements with propositional content p, and thereby commits to acting as though she believes.

Another verse by Meerabai, which holds trait of 'locutionary' speech act and can also be considered to be an 'expressive' is

My love is reserved for Girdhara Gopal
And for no one else. …
I left brothers and relatives
And all I possessed.…
I felt joy in the company of the devotees,
On beholding the world I wept.
I planted the creeper of love
And watered it with my tears.

(Alston 40)

It is not viable to be in the thick of the vachanas of the aforesaid sages and not come across the physicality of the words. To put it in a different way, poetry was devotional and metaphysical. At the heart of Bhakti was the desire to articulate. The tones of speech, the sharp mount and fall of the singer's voice and the recurrence of words and phrases often cannot be replicated. The literature of Bhakti draws on "magical – language", nurturing the impossibility, initiating a "secret language within a language" (Eliade 249). Candour of tone was preferred over eloquence. Diction was coarse, colloquial and simple. The poets created a theatre of participants, enthusiastic or disinclined. In most cases, the presence of listeners or spectators is only implied and that makes the' perlocutionary' speech act very effective. The poem implicated the listener into its world. The fierce confrontation, harassment, proselytizing, the vows and oaths were meant to draw the indolent and the reluctant into this world. The poets actively attempted to create a coterie of creative thinkers who would live by the questions posed by the poem. Most of them capitalized every kind of possibly usable figure of language to reach the insight of the listeners.

It is also worth revealing that the organized taxonomy and clarifications offered by linguists and philosophers of language have given edge to the already focused discussion. Therefore, to interpret the meaning correctly, role played by the situational framework has also been a reason for deliberation since language processing can be cogitated as an inter-mutual system in which any sort of incidental information can impact the style in which any other piece of information is formulated. It is in this situation, metaphor takes part in, since it transports what Richards call "a shift, a carrying over of a word from its normal use to a new use" (Constable, 178-179) on the basis of "an implied analogy in which one thing is imaginatively identified with another dissimilar thing" (Cuddon, 117). The importance of metaphor sprouted in Aristotle's Poetics who believed that: "the greatest thing by far is to be a master of metaphor. It is the 'one' thing that cannot be learnt from others; and it is also a sign of genius, since a good metaphor

implies an intuitive perception of the similarity in dissimilar". (Poetics, 1459 a) Metaphors are wide-ranging in their trait and Lakoff writes that "metaphor is fundamentally conceptual and not linguistic in nature" (1993: 244) In the present reach of this paper, the seminal work of Lakoff and Johnson has also been referred to because 'conceptual metaphor' theory came up with an aim to unearth organized mappings at the supra individual level, i.e. in our conceptual being (langue/competence). These mappings are then put into operation in definite contexts for the sake of communication between speakers at the individual level (parole/performance). In "Metaphors we live by", Lakoff and Johnson call attention to their theory by explaining that metaphors regulate our judgment and conduct, and that our conceptual system is metaphorical in its disposition. They brace the hypothesis with notions such as 'argument is battle'–mentioning that it is essential to see that we do not just discuss arguments in the terms of battle, but in reality, there are possibilities of vanquishing or losing arguments. Moreover, the authors surmise that the essence of metaphor is understanding and experiencing one kind of thing in terms of another. In the example 'argument is battle', we present 'argument' by means of something as serious as battle. The key idea of the 'conceptual' metaphor is the mapping of one domain (source) to another (target), as shown in the example 'argument is battle,' where 'argument' is the source and 'battle' the target field.

Lakoff and Johnson believe metaphorical concepts to be very well-heeled and complex which crop up from physical, cultural and diverse other experiences. A study of the select verses from Andal to Meera Bai reveals that the verses are not only elegant and simple but also very distinct and classic in the use of conceptual metaphors, encompassing all the three types —structural, oriental and ontological. In this context it is relevant to dig deep into the following piece of Anadal's vachana, also cited above, expressing her love for Vishnu

My surging breasts long to leap to the touch of His Hand
which hold aloft the flaming discus and the bright conch
Coax the world-measurer to caress my waist
To encircle the twin globes of my breasts

The expression "twin globes of my breasts" is a metaphor. It uses the word "globes" to compare the shape and roundness of breasts to the spherical objects known as globes. This figurative language adds a layer of imagery and could suggest fullness, roundness, or even a sense of wonder or beauty, thus advocating the use of conceptual metaphor.

Similarly in Akkamahadevi's vachanas, cited above, we find extensive use of metaphors, particularly to describe her intense devotion to Shiva (Channamallikarjuna) and to critique societal norms. She employs metaphors of illicit love, sexual union, and nature to express her spiritual journey and challenge conventional views on gender and devotion. Through the thick of expression of love for Chennamallikarjuna, when she was questioned regarding her nudity and the reason behind covering her body with her tresses, she answered fervidly through another vachana, which showcased the following lines

"If the fruit inside is not ripe, the
outer skin will not have its ripe colour;
If you see what causes sexual desire, you might feel hurt:
Thinking so, I have hidden it".

The entire segment "If the fruit inside is not ripe, the outer skin will not have its ripe colour; If you see what causes sexual desire, you might feel hurt, thinking so I have hidden it" contains metaphors. The first part uses the ripeness of a fruit as a metaphor for inner truth or essence, while the second part uses the concept of seeing the cause of sexual desire as a metaphor for understanding something deeply personal and potentially painful.

In another vachana, while expressing her heart wrenching cry for release from her familial bonds and be in harmony with Him, she recounts

"Like a silkworm weaving her house with love from her marrow,
and dying in her body's threads, winding tight, round
and round, I burn desiring what the heart desires".

The excerpt "like a silkworm weaving her house with love from her marrow, and dying in her body's threads winding tight, round and round," is rich with metaphors. The silkworm's act of creating a cocoon is used as a conceptual metaphor for human action and experience, specifically the act of creation. The phrase "love from her marrow" adds a layer of depth, suggesting the creation is not just a mechanical process but is fueled by deep passion and intrinsic motivation, "from her marrow" implying it comes from the very core of her being. The phrase "dying in her body's threads" is a powerful metaphor for the end of life, using the silkworm's death within its cocoon as a parallel to human mortality.

In Janabai's vachanas too, we find extensive use of metaphors. Particularly the phrases "I eat God, I drink God, I sleep on God, I buy God, I count God, I deal with God" are metaphorical in nature, suggesting the use of ontological metaphors. Ontological, because they allow us to understand abstract concepts by relating them to more concrete, physical actions. They employ actions associated with the physical world (eating, drinking, sleeping, buying, counting, dealing) to represent an unfathomable, all-encompassing spiritual bond with God. This entails a relationship where God is not merely an external being, but is incorporated into every aspect of the speaker's life and being.

Gangasati's vachanas also show evidences of metaphors as we see in the following extract, also cited earlier

"my head may separate from my body and roll off
but my torso will keep on fighting;
I have sacrificed my body"

The entire extract uses metaphor to convey the speaker's intense determination and resilience despite extreme physical or emotional hardship. The separation of the head (representing the mind or will) from the body (representing physical form) symbolizes a potential breaking point, yet the continued fighting of the torso (representing the core strength or spirit) signifies a steadfast pledge to persevere.

Meera's verses too were rich in metaphors of all kinds. Her transcendental thought fused with nobility of nature and her sublime love has produced poetry of the transcendent quality. To try to decipher such a realization casted into linguistic expression is not an easy task. Nonetheless, examples have been culled from a large plethora of Meera Bai's verses and classified in concordance with the three conceptual categories of metaphors provided by Lakoff and Johnson.

In the following few lines from the verse cited above Meera uses the metaphor of 'buy-sell' and thinks that her investment to buy 'the cherished' is actually nothing significant.

Sister, I have bought Govind in the market,
You only speak of it secretly,
But I proclaim it openly, beating a drum.
You say it was expensive,
But I say it was cheap,
And I measured it out on the scale

The lines in discussion also represent the reality that her body and life can be offered unquestionably to receive the love of Govind. The source domain is 'the act of shopping in the market' which has metaphorically been mapped to the target domain of 'offering oneself as a price' for achieving Shree Krishna's love. The example cited has been one of structural metaphors.

In the poem, "My Love is Reserved for Girdhara", Meera's linguistic expression has been catapulted by conceptual mapping from the concrete to the abstract especially perceptible in line-11-12 of the poem:

My love is reserved for Girdhara Gopal
And for no one else. …
I left brothers and relatives
And all I possessed.…
I felt joy in the company of the devotees,
On beholding the world I wept.
I planted the creeper of love
And watered it with my tears.

(Alston 40)

In the aforementioned verse, the causal sphere is 'creeper' (concrete) and the aim area is 'love' (abstract). The five characteristics listed by Lakoff and Johnson for ontological metaphors can be placed here: Referring – Girdhara Gopal (The Supreme One); enumerating – (leaving) brothers and relatives, possessions; locating Aspects: stipulation, Figuring out cause: company of the devotees; Setting Goals: planted the creeper; Motivating Actions –watered it with (my) tears.

The examples cited above, make it apparent that the idea of belonging controls a person's ontology: their analysis of the aspects of things, the points at which one particular thing becomes something else. It could be said to settle on a person's ecological responsiveness: their sagacity of understanding 'what belongs to me' and 'what is beyond'. Besides, ontological metaphors treat human states and emotion as containers. Herein, the notion of major and minor emotions like 'love-bliss' and 'love-whimpering-tears' create structures of familiarity, crafting a movement between ideas which help in making these lines exquisite examples of ontological metaphor.

Conclusion

Interpreting the Faith-driven Defiance of the Women Saints in Challenging the Hard-handed Hierarchy:

What can be said while wrapping up the work is that it is no epiphany, that even today, writers in Tamil, Kannada, Marathi and beyond, especially women, including those who do not write, attempt to draw their voyage into their own lives in order to make sense of the unspeakable. In the ancient traditions of poetry in India, meaning was regarded as a happenstance of form. The

traditions of magic and spiritual enquiry which created these poems were more interested in what the poem did than what it meant. The function of poetry was to project the 'yogin' into a particular state of awareness. The reason behind analyzing the language of the women mystics by deploying the speech acts is that speech acts affect the world by reshaping social norms. They alter what is permissible, obligatory, socially appropriate, transgressive, legal, and so forth. Most of the utterances of the sages that have been scrutinized for the present purpose have been construed to be defying the idiosyncrasy of women's speech that were sanctioned to be socially appropriate. It has been observed over the years by sociolinguists and theorists that there are certain very basic gender differences in the speech acts deployed in real conversations by men and women. It is observed that women have been more prone to questioning because of the constant uncertainty they endure and have been facing in their lives, for ages, while men are and have been in the habit of using more declarative statements. In refusal strategies too, women are strategically more indirect, than men as women engage in frequent acts of dissuasion, excuse, or apology. With regard to questions, women tend to use more question forms in conversation. Another very significant trait is that women use more linguistic hedges, such as "I think" and "sort of", which men do not. Women, also are found to use more paralinguistic responses, such as smiling and nodding. A further common belief about language is that women are more tentative in their speech. Lakoff (53-56) highlighted the use of the 'tag question' as an illustration of this claiming that women use more tag questions than men, who in turn by using them less, appear to be more assertive. Further research found that tag questions were more commonly associated with women's language (Siegler and Siegler 167-170), illustrating people's attitudes towards women's speech and its tentativeness. The study carried out for this purpose, however, found very different results with women displaying more dictatorial attitudes by engaging in conversations in the way that did not comply with the women's usual way of conducting. A deep probe into the lives and the language of the women saints has showcased the most daring mode of subversion by showing stern and unyielding attitude towards situations that created insurmountable social upheavals. Study of speech acts over the ages has proved that most of the speech acts have agent-relative authority conditions. That is, they can only be appropriately and successfully uttered by people with the right kind of social standing. Thus the sayings of these women sages, after being analyzed by employing speech act theory, do truly reflect upon the fact that women-saints, despite tremendous social constraints, did have the right kind of social standing and could manifest their longing to be in union with the God, even if the manner of manifestation was deeply amorous and sensual

The later part of the paper, where we have examined the vachanas by employing the framework of conceptual and ontological metaphor, suggested by George Lakoff, we have gained insights into the cultural assumptions and power dynamics that played a crucial role in our understanding of gender roles and identities. Lakoff's theory, in this context, has provided a powerful tool for analysis and for challenging detrimental gender stereotypes. By analyzing vachanas, we could unearth the fundamental cultural postulations and biases connected with gender roles and identities. The corpus based study on the metaphorical framing of women, which we focussed on, in the latter half of the paper, after a genuine effort of speech act analysis, has flamboyantly presented the faith-driven women as warriors, stubborn and unrelenting in their own ways to craft dominant spaces for themselves within a society that was predominantly drowned in the din of patriarchy.

Bibliography & References

Alston, A J. (1980). The Devotional Poems of Mirabai Translated with Introduction and Notes. Delhi. Motilal Banarsidass. (p. 60)

Austin, J. L. (2012). *How to Do Things with Words*. Beijing. Intellectual Property Publishing House.

Beauvoir, Simon de. (1988). *The Second Sex,* Eng. Tr. H. M. Parshley. London. Picador.

Cameron, Deborah. (1992) *Feminism and Linguistic Theory.* 2nd ed. London: Macmillan.

Constable, J. (2001). (Ed.) I.A. Richards and His Critics Vol:10. London. Routledge.

Cuddon, J.A. (1998). *The Penguin Dictionary of Literary Terms and Literary Theory.* Penguin Books.

Dehejia, Vidya (2008). *The Body Adorned: Sacred and Profane in Indian Art*. New York. Columbia University Press. (pp. 140-141).

Firth, J. R. (1957). *Papers in Linguistics 1934-51*. London. Oxford University Press.

Kolatkar, Arun (2010) *Collected Poems in English,* ed. Arvind Krishna Mehrotra, Northumberland: Bloodaxe Books.

Lakoff, G., and Johnson, M. (1980). *Metaphors we Live By.* Chicago and London. The University of Chicago Press.

Mulchandani, Sandhya.(2019). *For the Love of God: Women Poet Saints of the Bhakti Movement*. India. Penguin Random House.

Ramanujan. A.K. *Speaking of Shiva*. (1973). Translated with an Introduction by A. K, Ramanujan. England. Penguin Books.

Searle, J. R. (1964). What is a Speech Act? In A. P. Martinich (eds.). *The Philosophy of Language* (pp 221-239).

Shivaprakash. H. S. (2203). Vachanas of Akkamahadevi in *India International Centre*

Quarterly, vol. 30, no. 3/4, pp. 32–37, JSTOR, http://www.jstor.org/stable/23006121 . Accessed 20 Apr. 2023

Siegler, D and R. Siegler. (1976). "Stereotypes of Males' and Females' Speech". In *Psychological Reports 39*. (pp 167-70).

Venkatesh, Arundhati. (May 17, 2016). "Andal: The Tamil Female Saint Brought To You In A Very Satisfying Translation [Book Review]". In *Women's Web. Archived from* the original *on 26 January 2018*.

Wardhaugh, Ronald. (2002). *An introduction to sociolinguistics (4th ed.)*. Oxford. Blackwell Publishers Ltd.

Internet sources

www. shodhganga.inflibnet.ac.in

https://kcsujata.blogspot.com/2011/04/varanam-aayiram-by-sri-andal-verses-and.html

https://shaivam.org/to-practise/dont-postpone-the-grace/#gsc.tab=0

https://guftugu.indianculturalforum.in/category/gujarati/

Jadavpur Journal of Languages and Linguistics
ISSN: 2581-494X

Intergenerational Transmission of Tamil in Mauritius: A Sociolinguistic Analysis

Dahini Ponnen[1] *& S. Arulmozi*
University of Hyderabad, Telangana

ARTICLE INFO

Article history:
Received 23/09/2025
Accepted 13/10/2025

Keywords:
intergenerational transmission,
Mauritius,
Heritage language,
sociolinguistics,
language Shift,
multilingualism

ABSTRACT

This study investigates the intergenerational transmission of Tamil in Mauritius, focusing on how the language is passed down within Tamil-speaking families and supported by community institutions such as temples and Tamil schools. Despite Tamil's cultural and religious significance, its role in everyday communication is diminishing, particularly among younger generations. Using a qualitative sociolinguistic approach, this study draws on semi-structured interviews with three generations across twelve Tamil-speaking families, supplemented by observations in temples and educational settings. Thematic analysis reveals a complex interplay between emotional attachment to Tamil and decreased functional use. Parental ideologies, academic practices, and the dominance of Creole, English, and French in public life shape it. Findings suggest that while elders often retain fluency and use Tamil actively, parents exhibit partial transmission practices, and youth maintain only symbolic connections to the language. The study highlights the significance of family language policies, maternal influence, and institutional support in maintaining the Tamil language. It concludes by offering recommendations for community-based language planning that align with the realities of a multilingual, postcolonial society.

Introduction

Tamil is one of the longest-surviving classical languages and plays a significant role in shaping the identity of Tamil-speaking communities worldwide. Mauritius has been preserved primarily through religious practices and cultural traditions brought by indentured labourers from South India during the 19th century (Benedict, 2011). Over time, however, changes in the linguistic landscape, particularly the dominance of Creole as the national lingua franca and the institutional prominence of English and French, have undermined Tamil's everyday relevance (Rajah-Carrim, 2005).

Intergenerational language transmission is a critical determinant of linguistic vitality (Fishman, 1991). It encompasses the frequency of language use within families and the attitudes and beliefs that influence language choices. In multilingual and postcolonial societies like Mauritius, the pressure to conform to dominant linguistic norms often shifts from heritage languages to more socially and economically dominant ones (De Korne & Leonard, 2017).

This paper examines how Tamil is being transmitted across three generations in Mauritius. It examines the sociolinguistic dynamics that facilitate or hinder the transmission

[1] Corresponding Author Email ID: 22haph01@uohyd.ac.in & arulmozi@uohyd.ac.in

within home and community institutions. The study identifies challenges and opportunities for revitalising Tamil among younger generations.

Research Questions

This study is guided by the following research questions:

1. How is Tamil transmitted across generations in Mauritius, and what sociolinguistic factors facilitate or hinder its continuity?
2. What strategies and practices do Tamil-speaking families employ at home to maintain and transmit the Tamil language?
3. How do elders, parents, and youth differ in their proficiency, attitudes, and use of Tamil?
4. What role do Tamil schools and temples play in supporting or limiting intergenerational transmission?
5. How is Tamil perceived in relation to Creole, English, and French within Mauritian society?
6. How do emotional attachment and symbolic value influence actual language use?
7. What barriers prevent Tamil from being used as an everyday language of communication?
8. What opportunities exist for revitalisation, particularly through digital media, community initiatives, and institutional reform?

Literature Review

The study of intergenerational language transmission is central to understanding language shift and maintenance in multilingual societies. Fishman's (1991) foundational theory of Reversing Language Shift (RLS) highlights that for a minority language to survive, it must be passed on as a "mother tongue" within the family unit, which he designates as Stage 6 in the Graded Intergenerational Disruption Scale (GIDS). If children do not acquire the heritage language as their first language in the home, the shift process becomes almost inevitable. In such cases, symbolic affiliation to the language may remain, but active competence and everyday functionality steadily erode.

The concept of Family Language Policy (FLP) builds on Fishman's insight, offering a framework to analyse both explicit strategies and implicit patterns through which families manage multilingual practices (Spolsky, 2004; King, Fogle, & Logan-Terry, 2008). Parental ideologies, sociocultural expectations, socioeconomic status, and institutional language policies influence FLP. Curdt-Christiansen (2009) emphasises the "invisible language planning" performed by parents, particularly mothers, whose emotional investment and linguistic confidence can profoundly shape transmission outcomes. In this regard, maternal agency becomes a pivotal mechanism for sustaining linguistic heritage.

In Mauritius's diasporic and postcolonial context, Tamil has long existed within a dynamic linguistic hierarchy. Following the migration of South Indian Tamil labourers during the colonial period, Tamil was maintained through religious practices, community rituals, and oral transmission (Benedict, 2011). However, the rise of Mauritian Creole as the language of informal social interaction and the dominance of English and French in formal institutions have significantly marginalised Tamil (Rajah-Carrim, 2005). Even when Tamil is institutionally supported, through education or religious instruction, its usage is often restricted to ritualistic or literary functions (Mooneegadoo, 2018), which limits its relevance in modern communicative domains.

Eisenlohr (2006) offers an important extension to this understanding through his ethnographic work in Mauritius, arguing that Tamil functions within a system of "ritualised multilingualism" there, where language is preserved primarily as a symbol of sacred identity rather than as an everyday communicative tool. This notion complements Mooneegadoo's (2018) observation of Tamil's ritual confinement and foregrounds how symbolic attachment alone cannot guarantee linguistic continuity. Comparative studies from other Tamil diasporas reinforce this pattern. Research in Singapore, Malaysia, and Canada shows that Tamil retention is strongest when family practices, community schools, and state policies work in synergy to promote functional use rather than ceremonial learning (Annamalai, 2019; Rajendran, 2020). These models reveal that institutional modernisation, through conversational Tamil, project-based learning, and culturally relevant pedagogy, can transform heritage language education from a ritualistic to an interactive approach.

Studies from comparable multilingual settings reflect similar patterns. Pauwels (2016) notes that symbolic loyalty to a heritage language, while valuable, is insufficient to ensure intergenerational sustainability. Emotional attachment without functional reinforcement often leads to passive bilingualism or complete attrition. Additionally, Clyne and Kipp (1999) found, in their study of immigrant families in Australia, that the use of community languages decreased sharply among third-generation youth unless it was bolstered by consistent home use and meaningful interaction beyond ceremonial contexts.

In summary, the literature suggests a convergence of factors, including family ideologies, institutional reinforcement, and language prestige hierarchies, that collectively influence intergenerational transmission. In Mauritius, this complex interplay warrants a context-specific analysis to understand how Tamil survives, adapts, or retreats in contemporary family and community life.

A growing body of research highlights the role of digital media in reconfiguring heritage language use. Androutsopoulos (2015) and Sharma (2021) demonstrate how social platforms such as YouTube, TikTok, and Instagram allow youth to blend languages, perform identity, and revitalise minority tongues. Similar patterns are seen in other revitalisation movements: Welsh (Cunliffe, 2019), Hawaiian (Warner, 2018), and Māori (Johnson, 2020) communities have successfully leveraged online storytelling, podcasts, and memes to normalise heritage languages within popular culture. For Tamil, recent work by Kandasamy (2021) and Saravanan (2022) reveals that youth-led podcasts and devotional digital content contribute to a renewed sense of ownership over the language. This digital turn offers a new lens for examining how younger Mauritians incorporate Tamil into their social and online lives, an area this study aims to explore within the island's multilingual ecology.

Methodology

This study employed a qualitative, multi-method design combining semi-structured interviews and participant observation. A qualitative approach was chosen to capture the nuanced experiences, ideologies, and practices that shape intergenerational language transmission within Tamil-speaking families in Mauritius, where the complexity of multilingual repertoires often resists quantification alone (Braun & Clarke, 2006).

Sampling and Participants

Twelve Tamil-speaking families were selected through purposive sampling. The decision to work with twelve families was guided by the principle of theoretical saturation, a point at which

additional interviews no longer produced substantially new insights (Guest, Bunce, & Johnson, 2006). Families were chosen to reflect socio-economic, generational, and geographic diversity (urban and rural areas, public vs. private school backgrounds). Each family provided representatives across three generational cohorts: elders (60 years and older), parents (35–55 years old), and youth (13–30 years old). This cross-generational design allowed for comparative analysis of continuity, attrition, and transformation in language practices.

Semi-Structured Interviews

Interviews were conducted in Tamil, Creole, or English, depending on the participant's preference, to ensure comfort and authenticity of responses. The interview guide included questions on early language exposure, home communication practices, parental ideologies, literacy activities, code-switching norms, and attitudes toward Tamil in comparison with Creole, English, and French. Sample questions included: "What languages were spoken at home when you were a child?"; "Do you encourage your children to use Tamil at home, and if so, how?"; "How do you feel when Tamil is used in religious or social events?". Interviews typically lasted 45–60 minutes and were audio-recorded with consent.

Participant Observation

Observations were carried out in Tamil schools, temple classes, and cultural gatherings to complement the interview data. Particular attention was given to teaching methods, linguistic accessibility of rituals, and youth participation in institutional activities. Field notes documented both language use and the symbolic framing of Tamil in these settings.

Data Analysis

Interviews were transcribed verbatim and thematically analysed following Braun and Clarke's (2006) six-phase framework. Both inductive (emerging themes) and deductive (guided by research questions and theoretical models such as Family Language Policy and Fishman's GIDS) coding were applied. Codes were compared across generations and triangulated with observational data. NVivo software was used to organise and refine coding categories.

Ethical Considerations

Ethical clearance was obtained from the University of Hyderabad. All participants provided informed consent, and pseudonyms are used throughout to preserve confidentiality. Participation was voluntary, and participants were informed of their right to withdraw at any stage without consequence. Care was taken to ensure that sensitive reflections on family practices were handled respectfully and with cultural sensitivity.

Findings

Transmission Practices Across Generations

A clear generational divergence emerged in how Tamil is acquired, valued, and practised. Elders (60+) overwhelmingly reported Tamil as their first language, acquired naturally through immersive family and religious settings. Many noted that Tamil was the default language of household communication, temple activities, and early schooling experiences. One elder remarked, "We did not learn Tamil, we lived it."

In contrast, the parental generation (35–55 years) displayed a hybrid linguistic orientation. While some parents deliberately tried speaking Tamil at home, others defaulted to

Creole or English due to practical constraints or perceived futility. Notably, even parents who expressed pride in Tamil often admitted to inconsistent use of the language. A recurrent theme was a lack of confidence in their Tamil fluency, particularly in navigating modern topics. This led to a phenomenon of "passive transmission," where children hear Tamil but are rarely required to respond in it.

Youth participants (13–30 years old) reported minimal active engagement with Tamil outside of formal or religious contexts. Their exposure was typically limited to festivals, temple prayers, or Tamil language classes, and their proficiency ranged from basic comprehension to complete unfamiliarity with the language. Some youth expressed regret and a desire to improve their Tamil but cited a lack of practice opportunities, outdated teaching methods, and social irrelevance as barriers. One university student stated, "I wish I could speak Tamil fluently; it feels like a part of me. However, it is like having a piece of culture locked behind a wall."

Home Strategies for Language Maintenance

While intergenerational gaps in Tamil proficiency were evident, several families reported employing deliberate strategies to maintain Tamil within the home domain. Elders often employed oral traditions, such as storytelling, singing devotional songs, and reciting proverbs, to familiarise grandchildren with Tamil vocabulary and cultural idioms. In some households, parents introduced specific rules, such as requiring Tamil to be spoken during mealtimes or religious rituals. A few families encouraged literacy practices, including reading children's books in Tamil or practising script through temple handouts.

However, such strategies were inconsistently applied and frequently undermined by competing linguistic pressures. Code-switching norms often favoured Creole or English when discussing school-related topics or contemporary issues, as parents expressed limited confidence in conveying these ideas in Tamil. As a result, Tamil was often confined to symbolic or ritual contexts, producing what participants described as a "halfway transmission," exposure without enforcement. This illustrates the fragile balance between intentional language planning at the family level and the gravitational pull of dominant societal languages.

Language Ideologies and Symbolic Attachment

Instrumentalist rationales often shaped parental language ideologies. English was prioritised for educational and economic advancement, while Creole facilitated peer interaction and emotional bonding. Tamil, by contrast, was framed as a language of tradition, respect, and religion, worthy of reverence but difficult to integrate into contemporary life.

Among youth, Tamil was frequently described in symbolic or emotive terms: "sacred," "ancestral," "respected." However, this symbolic value did not correlate with fluency or usage. This mirrors Pauwels' (2016) findings on symbolic language maintenance, where heritage languages serve as identity markers rather than active communication mediums. Interestingly, some youth displayed signs of language guilt, feeling they had failed their ancestors or community by not learning Tamil properly.

Institutional Reinforcement and Its Limitations

Temples and Tamil schools remain central to the symbolic preservation of Tamil, yet their role in active transmission is constrained. Observations revealed that temple rituals emphasised classical chanting, scriptural recitation, and formalised prayers. While these reinforced Tamil's sacred status, younger participants often found the language inaccessible due to archaic vocabulary and

a lack of interactive explanation. Parents commented that temple spaces "preserve Tamil, but do not teach it," underlining their function as custodians of tradition rather than facilitators of contemporary language use.

Similarly, Tamil schools relied heavily on rote memorisation, grammatical drills, and literary passages. Several youth respondents characterised these classes as "boring," "irrelevant," or "too focused on exams," highlighting a pedagogical misalignment with their daily communicative needs. Teachers themselves often lacked training in second language pedagogy, defaulting instead to methods that privileged scriptural literacy over oral proficiency. This reflects not only pedagogical inertia but also policy-driven limitations, as curricula continue to frame Tamil as a classical subject rather than a living language. Without reform, these institutions risk reinforcing Tamil's symbolic role while failing to nurture communicative competence among youth.

Emotional Value versus Functional Use

Despite low usage, emotional attachment to Tamil was consistently high across all age groups. Tamil was a source of cultural identity, spiritual connection, and ancestral pride. However, the study revealed a significant disparity between emotional identification and functional fluency. This disconnect represents a significant challenge for revitalisation, as emotions alone do not lead to language retention unless supported by meaningful, everyday usage.

Several participants, especially the younger ones, expressed a desire for "relevant Tamil" - Tamil that could be used in casual conversation, on social media, or in contemporary storytelling. This highlights an unmet need: the creation of spaces where Tamil is not just preserved, but actively lived.

The data presented in this study illuminate a complex and evolving linguistic reality among Tamil-speaking families in Mauritius. At the heart of this reality lies a growing disjuncture between emotional affiliation and functional proficiency in Tamil, a phenomenon observed not only in Mauritius but in other diasporic and postcolonial communities globally (Pauwels, 2016; Clyne & Kipp, 1999). The findings suggest that the intergenerational transmission of Tamil is partial, symbolic, and domain-restricted, posing a substantial threat to the long-term vitality of the language.

Structural and Societal Barriers to Transmission

Beyond the challenges of home and institutional practice, broader structural and societal factors constrain Tamil's everyday use. The dominance of Mauritian Creole in informal interactions, coupled with the prestige of French and English in educational and professional spheres, significantly reduces the perceived utility of Tamil. Parents noted that success in school examinations and career advancement depended almost exclusively on proficiency in English and French, discouraging investment in Tamil beyond symbolic attachment.

Policy frameworks further marginalise Tamil, as heritage languages receive limited curricular time and are often treated as optional or extracurricular. Youth participants reported feeling stigmatised when speaking Tamil in public, reflecting wider societal ideologies that position Creole as inclusive and Tamil as sectarian or old-fashioned. This sense of embarrassment or fear of ridicule discourages active use, even among those with receptive competence.

Thus, barriers to Tamil transmission are not only linguistic but also ideological, institutional, and socio-political in nature. They reflect an ecology in which Tamil is structurally

disadvantaged, functioning more as a cultural emblem than as a practical medium of communication.

Discussion

The Role of Family Language Policy (FLP)

The study reinforces existing research on Family Language Policy (Spolsky, 2004; Curdt-Christiansen, 2009), highlighting the home as the primary site of linguistic negotiation. Households where Tamil was consistently spoken, involving active maternal participation, were significantly more successful in fostering language continuity. These cases exemplify what Curdt-Christiansen (2009) describes as "bottom-up language planning", where ideologies, intentions, and practices intersect to sustain minority languages.

However, most families in this study demonstrated a shift toward pragmatic multilingualism, with Tamil losing functional space to Creole and English. This pragmatic shift is often framed as a necessary adaptation to the realities of Mauritian society, where Tamil lacks instrumental utility. As in other multilingual contexts (De Korne & Leonard, 2017), parents tend to prioritise languages perceived as socially and economically empowering.

Language Shift and the Symbolic Trap

A recurring theme was the symbolic elevation of Tamil as sacred, ancestral, and respectable, even among youth who lacked proficiency in the language. While affirming identity, this symbolic status can also create what Pauwels (2016) terms a "symbolic trap", where heritage languages are revered but relegated to ceremonial and performative domains, with little incentive or infrastructure for actual usage. This finding aligns with Mooneegadoo's (2018) observation that Tamil is often "appreciated but not practised."

This dynamic poses a challenge for language revitalisation efforts. While community pride is essential, it is insufficient unless paired with opportunities for active engagement. Without this shift from symbolism to function, Tamil may continue its decline as a spoken language.

The Limits of Institutional Support

Although temples and Tamil schools contribute to the preservation of Tamil, their effectiveness in supporting intergenerational transmission is constrained by limited pedagogical innovation and narrow functional domains. Instructional practices rooted in rote learning and classical grammar fail to engage younger learners, especially those more familiar with digital, conversational, or hybrid modes of language (King & Fogle, 2013).

Moreover, while temple rituals provide cultural continuity, they rarely offer opportunities for interactive, real-world Tamil usage. This mirrors findings from Pauwels (2016), who argues that religious institutions can preserve form, but not necessarily function. Revitalisation efforts must consider how these spaces might evolve by incorporating bilingual explanations, youth-led initiatives, or informal Tamil discussion groups.

A Sociolinguistic Turning Point

The linguistic landscape suggests that Tamil in Mauritius is at a sociolinguistic crossroads. There is still time to shift trajectories if stakeholders act urgently and creatively. Digital media, community storytelling, and cross-generational dialogue represent underexplored opportunities

for Tamil revitalisation. Incorporating these into educational strategies could reposition Tamil as a heritage asset and a living, evolving language.

Revitalisation Pathways

Although Tamil in Mauritius faces significant challenges, the findings of this study also point toward opportunities for revitalisation that are both locally grounded and informed by comparative experiences. Participants themselves articulated a desire for "relevant Tamil," a version of the language that could be integrated into everyday interactions, digital spaces, and peer networks. Elders emphasised storytelling and devotional practices, while younger respondents expressed interest in conversational Tamil for use in social media, popular culture, and informal communication.

Family Level

At the family level, revitalisation must build on existing emotional attachment by translating it into habitual practice. Parents can play an instrumental role by introducing Tamil in daily interactions, such as mealtimes, shared reading, or family storytelling. Community groups could develop bilingual storybooks, parent–child reading sessions, or digital "Tamil Word of the Day" channels that make language learning informal yet consistent. Studies in family language policy confirm that frequent, low-pressure use at home is a key predictor of intergenerational continuity (Curdt-Christiansen, 2009; King & Fogle, 2013).

Institutional Level

At the institutional level, Tamil schools and temples must evolve from custodians of tradition to facilitators of communication. Drawing lessons from community schools in Singapore and Canada, curriculum design should combine conversational modules, role-play, and creative writing with traditional literacy and scripture (Annamalai, 2019; Rajendran, 2020). Temples can serve as hybrid learning spaces by providing bilingual explanations, youth-led discussion circles, and cultural workshops that integrate spoken Tamil into ritual contexts.

Societal and Digital Level

At the societal level, Tamil should be repositioned as a modern, hybrid language compatible with digital lifestyles. Young participants' interest in media and technology presents a natural pathway for engagement. Inspired by models such as Hawaiian, Māori, and Welsh revitalisation (Cunliffe, 2019; Warner, 2018; Johnson, 2020), Mauritian youth could develop Tamil podcasts, YouTube channels, memes, and online storytelling projects. These initiatives would not only modernise Tamil's image but also expand its functional domains. Policy reforms, such as allocating more curricular hours for Tamil, providing teacher training in communicative methods, and funding youth media initiatives, would further institutionalise these gains.

Conclusion

This study demonstrates that Tamil in Mauritius occupies a liminal yet resilient space, one that is deeply cherished yet restricted by structural and ideological constraints. Revitalisation, therefore, cannot rely solely on symbolic pride; it requires coordinated action across families, institutions, and digital communities. By embracing innovation while respecting tradition, Tamil can evolve as both a sacred heritage and a living language, ensuring its survival in a multilingual, postcolonial Mauritius.

References

Annamalai, E. (2019). *Language maintenance among Tamil diasporas: Continuity and adaptation. Journal of South Asian Languages and Linguistics, 6(2), 145–167*.

Androutsopoulos, J. (2015). Networked multilingualism: Practices of language and identity on social media. *International Journal of Bilingualism, 19*(2), 185–205.

Benedict, B. (2011). *Mauritius: Cultural Heritage and Identity.* University of Mauritius Press.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Clyne, M., & Kipp, S. (1999). *Pluricentric languages in an immigrant context: Spanish, Arabic and Chinese.* Mouton de Gruyter.

Crystal, D. (2000). *Language Death.* Cambridge University Press.

Cunliffe, D. (2019). Digital domains for minority languages: The case of Welsh. *Journal of Multilingual and Multicultural Development, 40*(5), 377–391.

Curdt-Christiansen, X. L. (2009). Visible and invisible language planning: Ideological factors in the family language policy of Chinese immigrant families in Quebec. *Language Policy, 8*(4), 351–375. https://doi.org/10.1007/s10993-009-9123-0

De Korne, H., & Leonard, W. Y. (2017). Indigenous language revitalisation. In O. Garcia, N. Flores, & M. Spotti (Eds.), *The Oxford Handbook of Language and Society* (pp. 727–746). Oxford University Press.

Eisenlohr, P. (2006). *Little India: Diaspora, Time, and Ethnolinguistic Belonging in Hindu Mauritius.* University of California Press.

Fishman, J. A. (1991). *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages.* Multilingual Matters.

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? *Field Methods, 18*(1), 59–82.

Johnson, H. (2020). *Māori and the Digital: Revitalisation and Resistance.* Te Herenga Waka Press.

Kandasamy, P. (2021). Devotion online: Tamil religious media and language continuity. *South Asian Diaspora, 13*(3), 241–258.

Kheirkhah, M., & Cekaite, A. (2018). Language maintenance in multilingual families: The role of family interaction and children's agency. *International Journal of Bilingualism, 22*(2), 153–170.

King, K. A., & Fogle, L. (2013). Family language policy and bilingual parenting. *Language*

Teaching, 46(2), 172–194. https://doi.org/10.1017/S0261444812000493

King, K. A., Fogle, L., & Logan-Terry, A. (2008). Family language policy. *Language and Linguistics Compass, 2*(5), 907–922. https://doi.org/10.1111/j.1749-818X.2008.00076.x

Mooneegadoo, B. (2018). *Language use and ethnic identity among Indo-Mauritians*. University of Mauritius.

Pauwels, A. (2016). *Language maintenance and shift.* Cambridge University Press. https://doi.org/10.1017/CBO9781139680445

Rajah-Carrim, A. (2005). Language use and attitudes among the Hindu community in Mauritius. *Journal of Multilingual and Multicultural Development, 26*(5), 394–411. https://doi.org/10.1080/01434630508669032

Rajendran, S. (2020). Tamil in Canada: Community schools and language planning. *International Journal of the Sociology of Language, 263*, 101–120.

Saravanan, V. (2022). Podcasting Tamil: Digital practices of heritage youth. *Language and Intercultural Communication, 22*(4), 456–472.

Sharma, B. (2021). *Heritage Languages in the Digital Age: Youth, Identity, and Media.* Routledge.

Spolsky, B. (2004). *Language Policy.* Cambridge University Press.

Warner, S. (2018). Hawaiian revitalisation in the 21st century: Challenges and innovations. *Language Documentation & Conservation, 12*, 1–23.

Jadavpur Journal of Languages and Linguistics
ISSN: 2581-494X

Cognitive Demands of Reading Devanagari Script in Formal Settings and Classrooms: A psycholinguistic perspective

Rupesh Kumar Pandey[1] *& Tariq Khan*[2]
[1]*University of Mysore*[1] *&* [2]*CIIL Mysore, Mysore*

ARTICLE INFO

Article history:
Received 13/08/2025
Accepted 01/10/2025

Keywords:
Devanagari script;
cognitive processes;
orthographic complexity;
script relativity;
Dyslexia

ABSTRACT

This study explores the cognitive complexities involved in processing the Devanagari script, i.e., an abugida script, a nonlinear orthographic system used for languages like Hindi, Marathi, Bodo, Konkani, Sanskrit, Nepali, and others. Devanagari's unique features, such as a 'schwa' like implicit vowel, its contextual deletion, and nonlinear vowel placements (matras), render cognitive challenges that are identifiably distinct from linear scripts like Roman and Latin. By reviewing the psycholinguistic literature, including studies on dyslexia and script relativity, this article examines how these script characteristics influence visual word recognition, phonological decoding, and syntactic parsing. The analysis highlights the increased cognitive load required to process Devanagari, especially for learners and individuals with reading difficulties such as alexia / dyslexia. The findings contribute to psycholinguistic theory by demonstrating how script complexity affects reading challenges and comprehension. Implications for educational practices, literacy programmes, and script design are discussed to enhance learning outcomes and script readability. This work underscores the necessity of script-specific approaches in multilingual / multiliterate / multiscript educational systems and supports the development of specific interventions for readers of complex scripts.

Introduction

Understanding the cognitive processes involved in reading is crucial in Psycholinguistics, the study of the psychological and neurobiological factors that enable humans to acquire, use, comprehend, and produce language (Leeper 1951). Cognitive processes refer to mental activities involved in acquiring knowledge and comprehension, including perception, memory, and reasoning (ibid.). The Devanagari script is one of the most widely used writing systems in the Indian subcontinent, used for several languages such as Hindi, Marathi, Nepali, and Sanskrit. Originating from the ancient Brahmi script, Devanagari exemplifies an abugida, a type of writing system in which consonants inherently carry a vowel sound that can be modified with diacritical marks (Bright 1999). Recognisable by its horizontal top line, shirorekha, connecting letters in a word, Devanagari has played a significant role in shaping cultural and linguistic identity within India (Ahmad 2008). Its development reflects historical efforts to unify the diverse linguistic communities of the region (P. T. Daniels 2008). In discussing Devanagari, it is important to distinguish among the related concepts of script, writing system, and orthography. A script is a set of symbols or characters used to represent the language visually; Devanagari serves as a

[1] Corresponding Author Email ID: rupeshpandey.ciil@gmail.com

script for multiple languages. A writing system consists of structural methods for arranging sounds into visual symbols which is demonstrated by the Devanagari abugida structure. Orthography defines the particular writing standards for a language including spelling rules and diacritical marks which differ among the languages written with Devanagari script. To appreciate the cultural and linguistic importance of Devanagari is to recognize these differences in clear terms. This is a step crucial in letter identification across various fonts of handwriting and requires attention to detail, as minute shape differences can alter letter identification.

Finally, symbol-sound mapping brings the identified shapes into contact with their corresponding sound (or meaning in logographic systems) to activate the left temporal lobe where phonology is processed (Price 2012). By repetition, readers automatise the process so that it can be performed quickly and efficiently, facilitating reading fluency. The Devanagari script's nonlinearity presents unique challenges to visual word recognition; the ability to read words accurately and quickly (Eddy, Grainger, Holcomb, & Gabrieli 2016), and to phonological decoding, sounding out written symbols to represent their auditory equivalents (Ziegler & Goswami 2005).

Devanagari script is a sophisticated and rich script in its culture and phonetic accuracy coupled with beauty. And yet, its sophistication, especially as far as conjunct consonants and variations between spoken and written language are concerned, can be challenging to learners and users alike. Variations between written and spoken language are universal to world scripts and hold true for the Devanagari script too. Redressing these shortcomings with simplified conventions or more recent technological advances can further increase its use and usability in today's environment.

In Devanagari, each consonant possesses a default vowel sound /ə/ unless altered by a vowel diacritic, negated by a halant mark, or suppressed by allographic representation. This default phonological feature necessitates the reader to inhibit or transform the default vowel constantly on the basis of contextual information, thus imposing a cognitive burden (Ray, Biswas, & Pal 2015). This implied vowel system necessitates higher phonological awareness and stricter working memory demands, according to Prakash, Rekha, Nigam, and Karanth (1993). The historical evolution of the Devanagari script and its orthographic rules has been influenced by linguistic and cultural forces, and these in turn affect the acquisition and processing of the script by learners (A. Daniels Belinda ; Sterzuk 2022).

Literature Review

This section provides a review of relevant literature while integrating significant research outcomes and theoretical perspectives. The study provides foundational context by examining available theoretical postulations alongside previous studies and primary field contributions.

The cognitive process of reading starts with visual processing, which involves brain recognition and interpretation of letter shapes and patterns (Dehaene et al. 2010; Rayner, Pollat- sek, Ashby, & Jr. 2012). The initial stage of reading requires readers to decode text by transforming visual symbols into corresponding sounds (C. Perfetti & and 2014). Understanding visual symbols in languages written with alphabets requires an essential decoding step. In reading development phonological processing involves readers matching letters to their corresponding sounds which serves as an essential step towards decoding skills and reading fluency (Share 1995). Semantic processing helps readers to understand words through their connections to existing knowledge and vocabulary. Semantic memory enables readers to understand text by allowing them to interpret the meaning of words and sentences (Cain & Oakhill 2006; C. A. Perfetti & Hart 2001). Readers perform syntactic processing to determine sentence structure and understand sentence

meaning. Understanding reading involves knowing how to combine words into sentences while also understanding how to read the words (Just & Carpenter 1992; Kintsch & van Dijk 1978). Furthermore, Working memory and comprehension monitoring play crucial roles in information retention during reading by verifying understanding (Cain 2009; Daneman & Carpenter 1980). Visual recognition combined with phonological decoding and semantic understanding along with syntactic parsing plus memory enables people to read and comprehend text through integrated cognitive processes.

Cognitive mechanisms engaged in reading Devanagari

The recognition of a sign or form as a letter during reading entails various cognitive processes that commence with visual perception. During this preliminary stage, the eyes perceive the visual characteristics of a symbol, its form, lines, curves, and orientation.The occipital brain handles this visual data and helps with basic shape recognition (Science News 2009). The visual word form area (VWFA) in the fusiform gyrus processes information by identifying shapes that appear as letters and words (McCandliss, Cohen, & Dehaene 2003). The VWFA plays a crucial role in identifying letters and words within scripts and connecting them to both alphabetic or logographic systems and sound knowledge. The next step involves finding the features after determining the shape. This step involves determining the specific characteristics that distinguish one letter from another. The Latin alphabet presents ‘b’ and ‘d’ as an example of distinct letter features. Subsequent to shape identification, feature detection is required. The ultimate stage of symbol-to-sound translation combines identified symbols with appropriate sounds or meanings in logographic scripts which activates phonological processing in the left temporal lobe (Price 2012). With consistent practice readers achieve automaticity that results in quicker and more accurate recognition of symbols which also improves their reading fluency. The nonlinear arrangement of the Devanagari script poses distinct challenges in visual word recognition , the capacity to accurately and swiftly identify written words (Eddy et al. 2016) , and phonological decoding, the mechanism of converting written symbols into their associated sounds (Ziegler & Goswami 2005).

Cognitive load and script complexity in Devanagari

Recent advancements in neurocognitive studies suggest that the complexity of writing scripts has a notable impact on the inherent cognitive load experienced by readers. Cai, Peng, and Ge (2024) explored the link between perceived cognitive strain and reading proficiency in Chinese, a logographic script with visual intricacies akin to those of Devanagari. An increased intrinsic load quickly led to lower comprehension scores unless learners engaged in targeted fluency building practices to enhance their ability to manage visual stimuli. The Orthographic Depth Hypothesis suggests that writing systems with intricate orthographic-phonological relationships, like Devanagari’s use of conjuncts and diacritics, make the decoding process more challenging. Moreover, according to the Psycholinguistic Grain Size Theory (Ziegler & Goswami 2005), students can handle complex information by aligning the processing of orthographic units, such as aksharas, with their phonological grain size, like syllables or morphemes, rather than focusing on individual graphemes.

Schwa deletion rules

Readers of the Devanagari script are likely to possess a good understanding of some phonological as well as morphological principles that regulate the circumstances under which the

/ə/ symbolized schwa has to be ignored or avoided when it comes to pronunciation. Principles such as have been established serve to go a long way towards ensuring that words are adequately represented and well decoded. Nevertheless, schwa deletion principles are bound to be usually very intricate and unpredictable due to the numerous variations that there are within different dialects and local forms of the language.

Nonlinear vowel placement

Matras, the vowel diacritics of the Devanagari script, are located in an unusual place in relation to consonants; they are written above, below, in front of, or even behind these consonants, resulting in a visually nonlinear arrangement (Bright 1996). This arrangement is a stark contrast to linear writing systems like the Latin script, where the letters are generally written in a simple sequential order. Study conducted by (Winskel & Perea 2014) have shown that reading nonlinear scripts engages more advanced visual processing strategies, which subsequently leads to longer fixation times and a higher frequency of regressions in reading. Moreover, advancements in computational methods, particularly in deep structure learning, have been harnessed to enhance the recognition of the Devanagari script, effectively addressing the challenges posed by its inherently nonlinear characteristics. (Susan & Malhotra 2020).

Impact on dyslexic readers

Dyslexia significantly hampers one's ability to process phonological information, and this issue becomes even more pronounced when dealing with scripts that have complex orthographies (Smythe, Everatt, & Salter 2004). Das, Banik, and Banik (2014) have demonstrated that dyslexic readers face considerable difficulties with the Devanagari script, largely due to its distinctive features, such as implicit vowels and its nonlinear structure.

Script relativity hypothesis

According to Winskel (2022), the structural features inherent in a specific script play a crucial role in shaping how we process information cognitively. Additionally, the Script relativity hypothesis suggests that when readers engage with different scripts, they develop unique cognitive strategies tailored to the specific demands and characteristics of their writing systems. For readers of Devanagari, this also involves a greater extent of spatial attention and context-dependent phonological decoding. Also, the 'native script effect' phenomenon lends support to the claim that the unique properties of a script have a very deep impact on the strategies of reading being used by readers (Gnanadesikan 2021).

Comparative studies

Comparative studies dealing with differences between Devanagari and Latin scripts list a variety of important differences in reading processes used by individuals. For example, the study conducted by Katz and Frost (1992) indicated that readers of Hebrew, using another nonlinear script with implicit vowels, have cognitive demands peculiarly similar to those of readers using the Devanagari script. The above observation lends strength to the hypothesis that orthographic complexity has universal effects on the reading strategies used by people. Furthermore, attempts at unification of several Indian scripts, e.g., the Bharatiya Lipi attempt, have been made to address the problem posed by script diversity. Such attempts can have far-reaching cognitive processing implications as well as literacy education implications as mentioned in (R. M. Jain & Kumar 2024). Moreover, the study of Sproat (2023b) traces the evolutionary history of symbols

and writing systems and hence derives valuable insights into how the complexity of various scripts has evolved and developed over the centuries.

Word (gloss)	Devanagari Script	Phonetic Form	Notes
Book	किताब	/kɪˈtaːb/	Inherent vowel suppressed in 'त'
Tomorrow	कल	/kəl/	Inherent vowel pronounced after 'क' and deleted after 'ल'
Time	वक्त	/ʋəkt/	Inherent vowel suppressed in 'क्' to form consonant cluster
House	घर	/gʱəɾ/	Inherent vowel pronounced after 'घ' and deleted after 'र'

Table 1. *Implicit Vowels in Devanagari*

Word (Gloss)	Devanagari Script	Intended Pronunciation	Literal Schwa Pronunciation	Notes
Time	वक्त	/ʋakt/	/ʋəkət/	Schwa deleted after penultimate and ultimate positions
Easy	सहज	/səɦəd͡ʒ//	/səɦəd͡ʒ/	Final schwa retained in literal form
Ram	राम	/ɾaːm/	/ɾaːm/	Common spoken omission of final schwa

Table 2. *Schwa Deletion in Devanagari*

Educational implications

The cognitive challenges linked to the Devanagari script carry important consequences for literacy teaching methods. Instructional approaches should incorporate strategies that reduce the increased cognitive burden, highlighting phonological awareness and visual processing skills as key areas (Nag & Snowling 2011). For those with dyslexia, it is vital that interventions are tailored to address the specific difficulties posed by this script (Karanth 2003). Additionally, the creation of Devanagari typefaces is crucial in improving readability for people with dyslexia, emphasizing the need for a thoughtful and intentional approach to script design (Lydon 2015).

Theoretical contributions to psycholinguistic processing of Devanagari Script

This section delves into fundamental theoretical frameworks that deepen our comprehension of the psycholinguistic mechanisms engaged in reading and interpreting the Devanagari script.

Cognitive strain from implicit vowels

The Working Memory Model (Baddeley & Hitch 1974) provides insights into the impact of processing implicit vowels on the phonological loop, which temporarily retains verbal information. The recurring necessity to consider these implicit vowels can deplete working memory resources rapidly, resulting in reduced cognitive efficiency during reading activities (Prakash et al. 1993).

Visual word recognition and nonlinear vowel placement

The Dual-Route Cascaded Model (Coltheart, Rastle, Perry, Langdon, & Ziegler 2001) proposes two pathways for recognizing words: a lexical route for familiar words and a sublexical route for decoding unfamiliar ones. In the case of Devanagari, the nonlinear placement of matras adds complexity to the sublexical route, forcing readers to piece together fragmented visual information. This observation aligns with findings from Sproat (2023b), which indicate that nonlinear scripts require more sophisticated visual processing strategies.

Script relativity and cognitive processing

This section of the article delves into how the process of visual interpretation influences our reading experiences, subsequently affecting how we create mental models of text and synthesize information. It examines how various writing systems, such as English, Chinese, Japanese, Korean, Thai, and Hindi/Devanagari, impact the reading experience, concentrating on distinctive features like letter forms, the orientation of text, and word spacing (Akamatsu 1999; Ben-Yehudah, Hirshorn, Simcox, Perfetti, & Fiez 2019; Das et al. 2014; Sun, Xu, Zhai, & Ji 2023; Winskel 2022). Regarding the relationship between language and thought, there are two principal viewpoints: linguistic relativism, which proposes that language shapes our cognitive processes, and linguistic determinism, which asserts that language outright dictates them. Interestingly, Whorf did not originally distinguish between these ideas; it was Roger Brown who clarified this differentiation after Whorf passed away in 1941. Studies in applied linguistics and psychology generally favor linguistic relativism, whereas linguistic determinism is less strongly supported. (Athanasopoulos, Samuel, & Bylund 2017).

Vowel Sound	Matra Symbol	Placement	Example Word	Notes
/i/	ि	Before the consonant	दिल	Matra placed before ‘द’
/u/	ु	Below the consonant	कुल	Matra placed below ‘क’
/e/	े	Above the consonant	सेब	Matra placed above ‘स’
/ai/	ै	Above the consonant	मैच	Matra placed above ‘म’

Table 3. *Nonlinear vowel placement*

Impact on dyslexic readers

Even if dyslexia presents a lifetime challenge, those with learning difficulties can efficiently manage its effects with suitable teaching approaches and tailored strategies. For those with dyslexia, processing phonological information is a major obstacle; this gets more difficult with sophisticated writing systems. Research indicates, for example, that dyslexic readers of Devanagari find great difficulty because of the script’s implicit vowels and nonlinear structure, complicating both phonological and visual handling. Devanagari introduces conjunct characters, vowel diacritics, and a horizontal line (shirorekha) which links letters within a word unlike alphabetic scripts like English, which have a simple linear sequence of letters, adding complexity to word segmentation and recognition. Moreover, implicit vowels, like the schwa deletion, have purposes in Hindi and Sanskrit that create additional challenges. Though in some cases the

schwa sound of the word (‘karnaa’) is dropped in its medial position, otherwise it has an inherent schwa sound. This erratic nature can be particularly difficult for dyslexic individuals especially since expecting these changes may result in reading errors

To mitigate these challenges, primary educators might implement strategies such as:

(1) Multisensory instructional techniques: Employing visual, auditory, and tactile methods to reinforce letter-sound connections and word recognition.
(2) Explicit phonological training: Clearly teaching schwa deletion rules and phoneme-grapheme correspondences to assist dyslexic learners in word decoding.
(3) Segmented and simplified text presentations: Displaying Devanagari words in a manner that visually separates complex conjuncts and diacritics to facilitate easier processing.
(4) Technological support: Using text-to-speech tools and phonetic keyboards that provide immediate feedback on pronunciation and spelling.
(5) Fonts Adjustments: Better reading and less visual confusion can be achieved by using fonts that are good for people with dyslexia and increasing the space between words and characters.

Teachers and researchers can reduce the challenges related to Devanagari’s complicated writing system by using these focused techniques, enabling reading to those with dyslexia.

Implications of the study in indian classrooms

The research emphasizes the necessity of modifying educational strategies in Indian classrooms to effectively tackle the distinct challenges associated with the Devanagari script. By prioritising the enhancement of phonological awareness and visual processing capabilities, educational curricula can be customized to provide students with a solid groundwork in literacy. This might involve integrating activities that reinforce sound-letter connections and visual identification of complex symbols. Additionally, employing pedagogical techniques that utilize technology, such as interactive digital learning tools, can boost student involvement and comprehension, ultimately resulting in better literacy achievements.

Educational strategies

To address the cognitive difficulties posed by Devanagari, our instructional strategies should focus on enhancing both visual processing skills and phonological awareness. Here are several program ideas that could be effective:

(1) Exercises for Phonemic Awareness: One should emphasise on the awareness and control of implicit vowels. Teachers can help their students determine whether the schwa sound is left out or pronounced by means of auditory discrimination activities (Karanth 2003).
(2) Multisensory Approaches: Combining visual, aural, and kinaesthetic techniques helps us to strengthen the link between matras and the related sounds. Students might, for example, track matras while vocalising the corresponding sounds (Nag & Snowling 2011).
(3) Technology Integration: Using instructional tools and applications created especially for learning the Devanagari script offers interactive and adaptive learning opportunities. Development of these tools can benefit much from advanced deep learning approaches for script recognition (Susan & Malhotra 2020).

Support for dyslexic readers

For those with dyslexia, the results of this study can truly make a difference since they direct the development of specific teaching plans and interventions. Understanding the special cognitive challenges the Devanagari script presents will help teachers create individualised literacy programs targeted at improving visual processing ability and phonological awareness. By means of multisensory teaching approaches combining visual, auditory, and kinaesthetic learning, dyslexic readers can more successfully link matras with their corresponding sounds (Wagner, Frankel, & Leighton 2023). For those with dyslexia, tech tools including software for script recognition driven by deep learning techniques can also provide customised support, making reading far more accessible.

Script and font design

By using simplified character forms that benefit early learners and people with dyslexia as well as by designing typefaces that stress matras to lower visual clutter (Sproat 2023a). Reading is improved by better script and font design. Furthermore improving cross-linguistic literacy could be harmonising elements among Indian scripts. Slow-motion audio and video displays help students, especially those who find reading challenging, better grasp of character shapes and stroke sequence. Institutions like the Central Hindi Directorate need to help to standardise the Devanagari script and lower variances during modernisation. The ideas of domain experts and educationists should also affect font design and script grammar so that innovations satisfy real language and educational requirements instead of being limited by technology constraints.

Writing systems result from thousands of years of reasoned debate, intellectual analysis, and cultural commitment. Practice, knowledge, and history have not been haphazard; they have rather deliberately generated every symbol and diacritical mark. Consider the historically recognised independent entities anusvāra and chandrabindu with particular prosodic and phonetic uses. In the same line, it provides not only shorthand, but also catches grammatical relationships and subtleties of omission that enhance Sanskrit knowledge. Born in the languages they teach, these are absolutely priceless intellectual jewels.

Furthermore, every element serves a different purpose and has historical background; hence, any attempt at standardisation should be handled carefully. Convenience helps us not to overlook variations in digital encodings or typesetting guidelines. These variances demand careful reading and professional corrections to maintain the whole expressive range of the script. Strong standardising procedures have to respect the conceptual integrity of these indicators and take geographical and technological variations into account.

Moreover, some marks are difficult to reproduce digitally, thus compromising crucial expressive tools when elements are combined or deleted just to simplify writing. Consolidation should only be taken into consideration in cases of actual duplication, in which two marks exactly provide information in all circumstances, and even then, it must be preceded by extensive academic review. In particular language or liturgical settings, each diacritical mark and conjunct is absolutely vital; without one, the script's adaptability may be inadvertently limited. Standardising norms will encourage growth rather than contraction. They have to clearly forbid the deletion of current elements and provide exact instructions for adding new symbols. Maintaining the subtle differences acquired over millennia helps one to adapt to the always changing needs of language. This approach guarantees that next generations can fully use the range of written language.

Standardisation should represent development rather than a return to reality from past. It must clearly outline a road of lifelong learning, welcome fresh ideas and technical developments, and really value the great legacy of the script. Millions of years of intelligent debate, intellectual analysis, and cultural commitment produce a writing system. Every sign and diacritical mark has been deliberately produced by experience, knowledge, and history. If we look at the historically identified separate entities: the ansvāra and the chandrabindu, they fulfil different phonetic and prosodic purposes. Likewise, it provides shorthand as well as marks grammatical links and subtle omissions that improve our knowledge of Sanskrit. Deeply ingrained in the languages they teach, these intellectual gems of great value are worth much.

Furthermore, every attempt at standardising should be handled carefully since every citizen has special importance and historical background. Convenience reminds us to see differences in digital encodings and typesetting standards. These modifications call for careful inspection and professional corrections to maintain the script's whole expressive capacity. Strong standardising rules have to respect the conceptual integrity of these markers and consider geographical and technological differences.

Standardisation of rules will encourage expansion rather than contraction. They especially have to forbid removing already-existing elements and offer precise directions for adding new symbols. Adapting to the always changing needs of language is mostly about maintaining those minute subtleties moulded over millennia. This strategy ensures that the range of written expression can be completely exploited by following generations. Standardising should also represent progress rather than a flight from reality. It must vividly show a road of lifelong learning, open fresh ideas and technological innovations, all the while respecting the rich legacy of the script. This means that the writing system remains strong and vivid, ready to satisfy the several needs of its users without losing the intricate meanings buried in every letter.

Enhancements in script and font design can significantly boost readability. Possible initiatives might include readability. Initiatives may involve:

- Use of Adaptive Font Technologies: Digital typesetting must be improved in terms of spacing and alignment, or adaptive font technology must be used to increase clarity without sacrificing script integrity. These methods handle readability in a more effective and balanced manner while also honoring the script's aesthetics.
- Simplified Character Shapes: There is a need to design educational materials that use simplified versions of Devanagari characters to aid recognition for early learners and individuals with dyslexia (Lydon 2015).
- Unifying Indian Scripts: Efforts should be made to create a unified script system that may simplify learning and improve literacy rates across different languages (V. Jain & Kumarasamy 2024).
- Creation of new technologies: Like slow-motion audio and video, new technologies should be developed for the display of electronic texts incrementally, enabling the learners, especially dyslexic ones, to grasp the shape of the characters and the stroke order involved in writing them.
- Institutional interventions: The Central Hindi Directorate (CHD) is a government institution that publishes the guidelines for the standardization of the Devanagari script and Hindi spelling. The role of this institution is of paramount significance as it can contribute towards achieving uniformity of design and reducing inconsistencies in spelling.

- Script change and modernization: Change in script is a modernisation process that is natural as well as mechanical. Irrespective of justification, some changes in scripts are quick to see, whereas some are noticeably slow. This creates the prevalence of variations. The script modernisation process should be conducted carefully so that the views of linguists and education givers are accommodated and inconsistencies are avoided.

Instructional practices to manage cognitive load

Thoughtfully crafted educational strategies can diminish unnecessary cognitive load by incorporating specific classroom activities.Truckenmiller and Tortorelli (2024) explored how oral reading fluency correlates with writing proficiency. They contended that employing activity oriented scaffolding techniques, such as guided prosody sessions alongside decoding exercises, enhanced students' reading and writing skills. Students improve their brain's ability to handle script complexities by practicing prosodic reading which fits the akshara-based Devanagari system and engaging in similar activities that focus on syllable-level processing. When students perform these exercises through culturally relevant stories it prompts activation of critical mental frameworks while combining cognitive load theory principles with culturally responsive teaching methods. These educational methods guide students from basic symbol recognition toward full mental and emotional engagement with content which improves automatic memory retrieval and understanding similar to systems that focus on syllables or morphemes rather than single graphemes.

Enhancement	Description	Expected Benefit
Colourful illustration of matras	Assign different colours to matras and consonants	Improves visual differentiation of components
Increased spacing	Add space between characters and matras	Reduces visual crowding and confusion
Simplified characters	Use basic shapes for complex letters	Facilitates easier recognition and learning

Table 4. *Font Design Enhancements for Devanagari*

Discussion

This work highlights several significant facets of script reading that demand further investigation. Reading from an unknown script is a cognitively taxing exercise that goes much beyond simple visual recognition, as demonstrated by a detailed examination of participant responses in various psycholinguistic studies. In addition to interacting with a script's visual form, readers must continuously try to decipher its structure, separate meaningful pieces, and connect them to their prior language knowledge. Readers have significant trouble decoding and understanding when word boundaries and familiar visual patterns are missing or distorted.

The use of segmentation and visual cues has been the most prevalent trend among participants, as per previous psycholinguistic studies. Many participants made an effort to visually deconstruct the script, attempting to distinguish between what appeared to be recognisable sequences, morphemes, or possible word units. Scripts with dense clusters of letters or no spacing were thought to be particularly challenging. This is consistent with earlier research that demonstrates the importance of segmentability, the capacity of readers to visually parse words or units, in reading fluency. The cognitive load rises when this segmentability is impaired, resulting in sluggish reading, hesitancy, or total understanding failure.

The study also highlighted the compensating mechanisms that readers use. To infer the content, readers refer to morphological intuition, sound approximations, or prior knowledge of similar-looking scripts. In many instances, readers try to use incomplete information to deduce the general meaning even when they are unable to read every sign. This kind of speculation demonstrates the reader's attempt to use top-down processing, in which interpretation is guided by expectations and contextual signals. However, the reading process becomes tedious and cognitively demanding when such tactics are not available because of a complete lack of familiarity with the script.

The necessity of providing script learners with tiered learning resources is a significant implication of this. Cognitive scaffolding can be achieved, for example, by teaching the script gradually, starting with solitary symbols and working your way up to compound forms and linked text. Additionally, students may be able to make connections between form and meaning if they are exposed to familiar context written in the new script, such as well-known words or phrases. The creation of script learning resources with integrated segmentation aids is one useful recommendation that comes from this study. Small spacing between morphemes or syntactic units during early exposure may help to ease the transition to fluent reading even if it is not customarily used in the script. Though they may not be on board with these developments, purists can really be useful teaching aids in the near future, reducing the first learning curve. A further crucial issue that arises in the analysis of reading is the necessity of consistent and clear scripts. It becomes a real difficulty for readers to interact with the text when its visual shapes are all over the place, thanks to eccentric fonts, small characters, or confusing ligatures. Therefore, particularly in the early phases of learning a script, it is imperative to choose techniques that are transparent, consistent, and favourable to education. Until students have a strong foundation of knowledge, we should not bombard them with sophisticated ligature systems or elegant variants.

Word (Gloss)	Devanagari Script	Intended Pronunciation	Literal Pronunciation (with Schwa)	Notes
Time	वक्त	/ʊakt/	/ʊəkət/	Schwa deleted after penultimate and ultimate positions
Easy	सहज	/səɦəd͡ʒ/	/səɦəd͡ʒ/	Schwa deleted after ultimate position
Ram	राम	/raːm/	/ɾaːm/	Schwa deleted after ultimate position

Table 5. *Schwa deletion in Devanagari (word-final schwa deletions)*

Moreover, including some auditory support will greatly improve the reading experience. Learners will find it simpler to link those visual shapes with known phonological forms if we combine visual script input with sound cues, such as phonemic breakdowns or pronunciations. This kind of multimodal reinforcement can reduce the guesswork and help readers better remember the shapes of the script, especially when they are encountering it for the first time.

The necessity of task simplicity and topic familiarity in the early phases of script exposure is another finding from this study. The combined cognitive strain can be too much for readers to handle when they are presented with both unfamiliar content and an unfamiliar script. As a result, one of the most effective ways to introduce readers to script reading is to use terms they

are already familiar with or sentence structures they can rely on in the new script. The reader can concentrate on script decoding without having to figure out new lexical or syntactic information. Also, several words serve as iconised forms in script processing, which means that instead of being decoded phonetically step-by-step, they are recognised and understood as cohesive visual or symbolic wholes. Hindi words including 'ऋषि' (rishi), 'ॐ' (om), 'स्वास्तिक' (swastik), and 'त्रिशूल' (trishool). Though at first, they seem complicated, these words are rather easy to understand. Our consistency of exposure and cultural awareness help us become familiar. Rather than spending much time in graphemic analysis, we usually identify these words from their spelling and visual cues, which set off an instantaneous cognitive connection. English speakers choose difficult words like Wednesday or colonel, but Hindi readers rely on thorough knowledge for culturally relevant or aesthetically pleasing words. This emphasises the need of teaching literacy using both symbolic recognition and phonetic decoding. By linking these known forms with stories, contextual images, or useful applications, teachers can improve script fluency and retention. In scripts rich in cultural or symbolic significance, a combined approach combining picture recognition with decoding can greatly help one to acquire complex or irregular forms.

This work also opens fresh directions for next studies on script acquisition and reading techniques. Although the present work concentrated on qualitative comments and observational data, future studies should investigate quantitative measures including reading speed, accuracy rates, and error patterns over many runs. Controlled tests help to further validate the visual components, space, ligatures, and character complexity that most affect reading fluency in new scripts.

Furthermore, this study can form the foundation for developing teaching materials for multilingual and multiscript students. It is increasingly important for interactive tools allowing students to practice reading unknown scripts in a controlled and flexible manner to be available. Among these reading tools could be segmentation cues, audio help, and a progressive complexity increase. In multilingual contexts, especially those where many scripts are actively or simultaneously used, such developments could improve script literacy and learning outcomes (South Asia, for example).

Another path of future research could be looking at how individual differences influence script learning. Although some readers might rely more on aural signals or verbal intuition, others might have better visual memory. Finding these learner profiles could enable the creation of customised script learning approaches catered to individual capabilities. Moreover, looking at how past knowledge of other writing systems, such as alphabetic, abugida, and logographic affects script adaptation will help us to better grasp human reading.

Matra	Placement	Challenge	Example	Notes
ि (/i/)	Before	Visual processing difficulty as matra appears before but is pronounced after	किताब	Matra 'ि' appears before 'क', but pronunciation follows 'क'
ा (/a:/)	After	Consistent placement, less challenging	राम	Matra 'ा' follows 'र', pronounced as /ra:m/
ौ (/au/)	Above	Requires upward eye movement, interrupts linear reading flow	कौशल	Matra 'ौ' placed above 'क', modifying vowel sound

ु (/u/)	Below	Requires downward eye movement, adds visual complexity	गुरु	Matra 'ु' below 'ग', pronounced as /guru:/

Table 6. *Impact of nonlinear matra placement on reading in Devanagari*

At last, this can affect script policy, education, and literacy planning. The research shows that ignorance of a script can impede understanding even in cases when the language is known. Access to script learning becomes vital for egalitarian education in multilingual communities where several scripts coexist. Therefore, by designing inclusive literacy policies considering reader readiness and script accessibility, one can close the difference between oral competency and reading proficiency.

Our research concludes that, particularly for unknown scripts, reading scripts is a time-consuming, strategic, and even frustrating task. However, the same trends, namely, reader preferences, common issues, and strategies, also point to clear solutions. Improving visual segmentation, simplifying script form, offering scaffolded learning, and integrating multimodal assistance are all workable ways to make script reading easier. By building on these findings, future research and teaching can greatly progress a more accessible and readable script acquisition method.

Recommendations

Devanagari has evolved into the abugida script. It is now via centuries of scholarly refinements. Each symbol or diacritic mark representing consonants and vowels has a significant cultural and phonological value. Its sheer richness, however, implies that schwas that organically disappear during reading and matras (that float above, below, or next to consonants) have a cognitive burden on students on the one hand. They are schemes/processes to encode or represent linguistic information, on the other hand. Therefore, due care is required while incorporating modifications. A reform that ignores these aspects runs the risk of not only becoming ineffective or cosmetic but also of adding to the cognitive load during the reading process.

In light of these intrinsic requirements, explicit schwa-deletion training must be prioritised in classroom education rather than being relegated to a footnote. Due to the inherent unpredictability of the /ə/ sound, rigorous drilling and contextual exercises are necessary; otherwise, students are left in a guessing game that erodes their confidence and fluency. According to scholarly opinion, focused schwa instruction is essential to understanding Devanagari's phonological logic and is not optional.

Similarly, it is no longer acceptable to rely exclusively on conventional, visual-only teaching techniques. When matras interfere with linear visual flow, the difficulties faced by phonologically processing and dyslexic learners are exacerbated. Multisensory methods, which combine visual highlighting, spoken reinforcement, and tactile tracing, are tried-and-true scaffolding rather than educational indulgences. A sizable portion of pupils would be denied the opportunity to achieve if these resources were removed.

These problems are exacerbated by the way Devanagari text is presented visually. Uniform spacing creates crowding that slows down quick word recognition, and dense, static fonts obscure important differences between conjuncts and vowel marks. Adaptive typefaces with unambiguous glyph designs and kerning that can be adjusted must be developed and adopted. It is also reasonable to include tiny segmentation cues in beginner materials since they direct the viewer's attention without changing tradition.

When it comes to policy, every single attempt that is made toward standardization or script extension needs to be meticulously supervised by a permanent committee that is made up of educators, linguists, and typographers. Conjunct rules and diacritical markings that are already in use should never be abandoned with the intention of making things more convenient. On the other hand, the introduction of extra symbols to assist unwritten or minority languages is not only vital but also legitimate. However, this process must be driven by the same broad consultation that was used to develop the foundation of Devanagari.

It is impossible for any of these crucial reforms to be maintained without the commitment of the institutions. The cognitive and cultural features of Devanagari should be incorporated into each and every certification course that is offered by teacher-training programs. This would ensure that educators receive a comprehensive understanding of not only how to teach the script but also the relevance of its intricacies. In the absence of such mandates, even the most effective tools and procedures would continue to be underutilised.

In the end, theoretical discussions have to give way to the rigidity of actual research. We owe it to the students as well as to the heritage of the script to conduct controlled studies that evaluate reading speed, comprehension, and error patterns both before and after each intervention. In order to confirm or challenge our proposals, only systematic evidence is required. This ensures that the growth of Devanagari is based on measured progress rather than straightforward progress.

Conclusion

The unique difficulties presented by the Devanagari script, which features implicit and nonlinear vowel placements, are examined in this paper. These characteristics require readers to use particular decoding techniques and activate extra mental resources, so requiring more cognitive effort. Understanding these needs is crucial, especially in literacy and educational environments where specific strategies can help readers, including those with dyslexia, very much. Considering the special qualities of the script helps us to improve psycholinguistic models and teaching techniques, so helping people in multilingual surroundings. Developing good teaching strategies and designing customized interventions depend on an awareness of these complications.

References

Ahmad, R. (2008). Scripting a new identity: The battle for Devanagari in nineteenth century India. *Journal of Pragmatics*, *40*(7), 1163-1183. Retrieved from www.sciencedirect.com/science/ article/pii/S0378216607001051

Akamatsu, N. (1999, August). The effects of first language orthographic features on word recognition processing in English as a second language. *Reading and Writing*, *11*(4), 381–403. Retrieved from www.doi.org/10.1023/A:1008053520326

Athanasopoulos, P., Samuel, S., & Bylund, E. (2017, 01). The psychological reality of spatio-temporal metaphors.. Retrieved from www.doi.org/10.1075/hcp.56.12ath

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), (Vol. 8, p. 4789). Academic Press. Retrieved from www.sciencedirect.com/science/article/pii/S0079742108604521

Ben-Yehudah, G., Hirshorn, E. A., Simcox, T., Perfetti, C. A., & Fiez, J. A. (2019). Chinese English bilinguals transfer l1 lexical reading procedures and holistic orthographic coding to l2 english. *Journal of Neurolinguistics*, *50*, 136-148. Retrieved from www.sciencedirect.com/science/article/pii/S0911604417300921(Cross-linguistic

perspectives on second language reading)
Bright, W. (1996). Devanagari. In P. T. Daniels & W. Bright (Eds.), *The world's writing systems* (pp.
384–390). Oxford University Press.
Bright, W. (1999). A matter of typology: Alphasyllabaries and abugidas. *Written Language & Literacy*, *2*(1), 45–55.
Cai, Y., Peng, X., & Ge, Q. (2024). The tango between perceived cognitive load and enjoyment of reading in determining reading achievement. *Reading and Writing*, *37*, 703–729. Retrieved from www.doi.org/10.1007/s11145-024-10621-9
Cain, K. (2009). Children's reading comprehension difficulties. *Contemporary Perspectives on Reading and Spelling*, 59–75.
Cain, K., & Oakhill, J. (2006). Profiles of children with specific reading comprehension difficulties.
British Journal of Educational Psychology, *76*(4), 683–696.
Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). Drc: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204.
Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450-466. Retrieved from www.sciencedirect.com/science/article/pii/S0022537180903126
Daniels, A., Belinda ; Sterzuk. (2022). Indigenous language revitalization and applied linguistics: Conceptualizing an ethical space of engagement between academic fields. *Canadian Journal of Applied Linguistics / Revue canadienne de linguistique appliquée*, *25*(1), 1–18. Retrieved from www.doi.org/www.doi.org/10.37213/cjal.2022.31841
Daniels, P. T. (2008). 14 writing systems of major and minor languages. In *Language in South Asia*
(pp. 285–301). Cambridge University Press.
Das, T., Banik, A., & Banik, S. (2014). Dyslexia in the Bengali language. *Annals of Dyslexia*, *64*(1), 1–17.
Dehaene, S., Nakamura, K., Jobert, A., Kuroki, C., Ogawa, S., & Cohen, L. (2010). Why do children make mirror errors in reading? neural correlates of mirror invariance in the visual word form area. *Neuroimage*, *49*(2), 1837–1848.
Eddy, M., Grainger, J., Holcomb, P., & Gabrieli, J. (2016, 09). Orthographic and phonological processing in developing readers revealed by erps. *Psychophysiology*, *53*. doi:
Gnanadesikan, A. E. (2021). S1: The native script effect. In Y. Haralambous (Ed.), *Grapholinguistics in the 21st century, 2020. proceedings* (Vol. 4, pp. 103–123). Brest, France: Fluxus Editions. Retrieved from www.academia.edu/48900948/S1_The_Native_Script_Effect
Jain, R. M., & Kumar, T. (2024). Bharatiya lipi: Unifying indian scripts. In B. Alareeni & A. Hamdan (Eds.), *Navigating the technological tide: The evolution and challenges of business model innovation* (pp. 72–81). Cham: Springer Nature Switzerland.
Jain, V., & Kumarasamy, D. (2024). Factors driving India's growth: Challenges and policy measures.*The Economists' Voice*.
Just, M. A., & Carpenter, P. A. (1992, January). A capacity theory of comprehension:

Individual differences in working memory. *Psychological Review*, *99*(1), 122–149. doi:

Karanth, P. (2003). *Cross-linguistic study of acquired reading disorders* (1st ed., Vol. 1). New York: Kluwer Academic/Plenum Publishers. Retrieved from www.doi.org/10.1007/978-1-4419-8923-9 (Springer Book Archive, Published: 31 Dec 2003 (hardcover), 17 Sep 2012 (softcover), 06 Dec 2012 (eBook))

Katz, L., & Frost, R. (1992). Chapter 4 the reading process is different for different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (Vol. 94, p. 67-84). North-Holland. Retrieved from www.sciencedirect.com/science/article/pii/S0166411508627892

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*(5), 363–394. doi:Leeper, R. (1951). Cognitive processes.

Lydon, K. S. (2015). *Designing Devanagari type: The effect of technological restrictions on current practice* (Doctoral dissertation). Iceland Academy of the Arts

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, *7*(7), 293–299.

Nag, S., & Snowling, M. J. (2011, July). Cognitive profiles of poor readers of kannada. *Reading and Writing*, *24*(6), 657–676. Retrieved from www.doi.org/10.1007/s11145-010-9258-7

Perfetti, C., & and, J. S. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*(1), 22–37. Retrieved from www.doi.org/10.1080/10888438.2013.827687

Perfetti, C. A., & Hart, L. (2001). The lexical basis of comprehension skill.

Prakash, P., Rekha, D., Nigam, R., & Karanth, P. (1993). Phonological awareness, orthography, and literacy. In R. J. Scholes (Ed.), *Literacy and language analysis* (pp. 55–70). Hillsdale, NJ: Lawrence Erlbaum Associates. Retrieved from www.academia.edu/5436026/Phonological_Awareness_Orthography_and_Literacy_by_P_Prakash_D_Rekha_R_Nigam_and_P_Karanth

Price, C. J. (2012). A review and synthesis of the first 20years of pet and fmri studies of heard speech, spoken language and reading. *NeuroImage*, *62*(2), 816-847. Retrieved from www.sciencedirect.com/science/article/pii/S1053811912004703 (20 YEARS OF fMRI) doi:

Ray, A., Biswas, A., & Pal, U. (2015). Schwa deletion in devanagari script and its implication in offline handwriting recognition. In *Proceedings of the 13th international conference on document analysis and recognition (icdar)* (pp. 296–300).

Rayner, K., Pollatsek, A., Ashby, J., & Jr., C. C. (2012). *Psychology of reading* (2nd ed.). New York: Psychology Press. Retrieved from www.doi.org/10.4324/9780203155158 doi:

Science News. (2009, November 20). Reading in the brain: The science and evolution of a human invention by stanislas dehaene. *Science News*. Retrieved from www.sciencenews.org/article/reading-brain-science-and-evolution-human-invention-stanislas-dehaene (Accessed: 2025-05-28)

Share, D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition*, *55*(2), 151–218. Retrieved from www.doi.org/10.1016/0010-0277(94)00645-2 (discussion 219–226)

Smythe, I., Everatt, J., & Salter, R. (Eds.). (2004). *International book of dyslexia: A*

crosslanguage comparison and practice guide. Chichester, UK: John Wiley & Sons. Retrieved from www.wiley.com/en-gb/International+Book+of+Dyslexia%3A+A+Cross+Language+Comparison+and+Practice+Guide-p-x000211636

Sproat, R. (2023a). Simulating the evolution of writing. In *Symbols: An evolutionary history from the stone age to the future* (pp. 145–178). Cham: Springer Nature Switzerland. Retrieved from www.doi.org/10.1007/978-3-031-26809-0_7

Sproat, R. (2023b). Writing systems. In *Symbols: An evolutionary history from the stone age to the future* (pp. 91–112). Springer. doi:

Sun, C., Xu, T., Zhai, C., & Ji, H. (2023). *Incorporating task-specific concept knowledge into script learning*. Retrieved from arxiv.org/abs/2209.00068

Susan, S., & Malhotra, J. (2020, 11). Recognising devanagari script by deep structure learning of image quadrants. *DESIDOC Journal of Library Information Technology*, *40*, 268-271. doi:

Truckenmiller, A. J., & Tortorelli, L. S. (2024). Automaticity in writing in response to reading: Relations between oral reading fluency and compositional writing fluency in grades 3–5. *Reading & Writing Quarterly*. (Advance online publication)

Wagner, C. J., Frankel, K. K., & Leighton, C. M. (Eds.). (2023). *Becoming readers and writers: Literate identities across childhood and adolescence*. Taylor & Francis.

Winskel, H. (2022). Script relativity hypothesis: evidence from reading with different spatial layouts and varied lexical tone. *Reading and Writing*, *35*(6), 1323–1341. Retrieved from www.doi.org/10.1007/s11145-021-10225-7

Winskel, H., & Perea, M. (2014). Can parafoveal-on-foveal effects be obtained when reading an unspaced alphasyllabic script (thai)? *Writing Systems Research*, *6*(1), 94–104. Retrieved from www.doi.org/10.1080/17586801.2013.843440

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, *131*(1), 3–29. Retrieved from www.doi.org/10.1037/0033-2909.131.1.3